Media Space

an analysis of spatial practices in planar pictorial media

a thesis submitted to Middlesex University
in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

Stephen Boyd Davis
Lansdown Centre for Electronic Arts
Middlesex University

June 2002

Abstract

The thesis analyses the visual space displayed in pictures, film, television and digital interactive media. The argument is developed that depictions are informed by the objectives of the artefact as much as by any simple visual correspondence to the observed world. The simple concept of 'realism' is therefore anatomised and a more pragmatic theory proposed which resolves some of the traditional controversies concerning the relation between depiction and vision. This is then applied to the special problems of digital interactive media.

An introductory chapter outlines the topic area and the main argument and provides an initial definition of terms. To provide a foundation for the ensuing arguments, a brief account is given of two existing and contrasted approaches to the notion of space: that of perception science which gives priority to acultural aspects, and that of visual culture which emphasises aspects which are culturally contingent.

An existing approach to spatial perception (that of JJ Gibson originating in the 1940s and 50s) is applied to spatial depiction in order to explore the differences between seeing and picturing, and also to emphasise the many different cues for spatial perception beyond those commonly considered (such as binocularity and linear perspective). At this stage a simple framework of depiction is introduced which identifies five components or phases: the *objectives* of the picture, the *idea* chosen to embody the objectives, the *model* (essentially, the visual 'subject matter'), the characteristics of the *view* and finally the substantive *picture* or depiction itself. This framework draws attention to the way in which each of the five phases presents an opportunity for decision-making about representation. The framework is used and refined throughout the thesis.

Since pictures are considered in some everyday sense to be 'realistic' (otherwise, in terms of this thesis, they would not count as depictions), the nature of realism is considered at some length. The apparently unitary concept is broken down into several different types of realism and it is argued that, like the different spatial cues, each lends itself to particular objectives intended for the artefact. From these several types, two approaches to realism are identified, one prioritising the creation of a true illusion (that the picture is in fact a scene) and the other (of which there are innumerably more examples both across cultures and over historical time) one which *evokes* aspects of vision without aiming to exactly imitate the optical stimulus of the scene. Various reasons for the latter approach, and the variety of spatial practices to which it leads, are discussed. In addition to analysing traditional pictures, computer graphics images are discussed in conjunction with the claims for realism offered by their authors. In the process, informational and affective aspects of picture-making are distinguished, a distinction which it is argued is useful and too seldom made.

Discussion of still pictures identifies the evocation of movement (and other aspects of time) as one of the principal motives for departing from attempts at straightforward

optical matching. The discussion proceeds to the subject of film where, perhaps surprisingly now that the depiction of movement is possible, the lack of straightforward imitation of the optical is noteworthy again. This is especially true of the relationship between shots rather than within them; the reasons for this are analysed. This reinforces the argument that the spatial form of the fiction film, like that of other kinds of depiction, arises from its objectives, presenting realism once again as a contingent concept.

The separation of depiction into two broad classes – one which aims to negate its own mediation, to seem transparent to what it depicts, and one which presents the fact of depiction ostensively to the viewer – is carried through from still pictures, via film, into a discussion of factual television and finally of digital interactive media. The example of factual television is chosen to emphasise how, despite the similarities between the *technologies* of film and television, spatial practices within some television genres contrast strongly with those of the mainstream fiction film. By considering historic examples, it is shown that many of the spatial practices now familiar in factual television were gradually expunged from the classical film when the latter became centred on the concerns of narrative fiction.

By situating the spaces of interactive media in the context of other kinds of pictorial space, questions are addressed concerning the transferability of spatial usages from traditional media to those which are interactive. During the thesis the spatial practices of still-picture-making, film and television are characterised as 'mature' and 'expressive' (terms which are defined in the text). By contrast the spatial practices of digital interactive media are seen to be immature and inexpressive. It is argued that this is to some degree inevitable given the context in which interactive media artefacts are made and experienced – the lack of a shared 'language' or languages in any new media. Some of the difficult spatial problems which digital interactive media need to overcome are identified, especially where, as is currently normal, interaction is based on the relation between a pointer and visible objects within a depiction. The range of existing practice in digital interactive media is classified in a seven-part taxonomy, which again makes use of the *objective-idea-model-view-picture* framework, and again draws out the difference between self-concealing approaches to depiction and those which offer awareness of depiction as a significant component of the experience. The analysis indicates promising lines of enquiry for the future and emphasises the need for further innovation. Finally the main arguments are summarised and the thesis concludes with a short discussion of the implications for design arising from the key concepts identified – expressivity and maturity, pragmatism and realism.

Dedication

This thesis is dedicated to the memory of John Lansdown, former leader of the Centre for Electronic Arts at Middlesex University now named in his honour, who until his death was the Director of Studies for this research and was a source of encouragement, advice and inspiration.

Acknowledgements

I gratefully acknowledge the advice and support of my supervisors, Professor Huw Jones and Gordon Davies of the Lansdown Centre and Graham Howard of Art of Memory. Gordon Davies and Magnus Moar took on additional teaching and administration which released me to undertake my research.

I am grateful to many for their comments on specialist subjects, in particular Professor Roy Armes for his help in the analysis of film including the loan of his own books. I was fortunate in having access to pre-publication material by Jesse Norman of University College London, and personal communication from Professor Richard Gregory, Professor Julian Hochberg and Dr Jeroen Goossens in clarification of various points. I was greatly helped by Fred Gill, technical director of Kaboom Studios, and his staff, who were generous with time and information. The comments of my external examiners, Professor Clive Richards of Coventry University and Professor Gordon Clapworthy of DeMontfort University, were extremely helpful in refining this final version of the thesis document.

I thank my wife Vanessa and son Edward for tolerating the research and writing which at times took over all our lives, and thank them both for their practical help of many kinds.

Contents

1	Introduction	1
2	Approaches to the study of space	10
3	Depth perception and depiction	20
4	Spatiality and realism	55
5	Screen space I: the spatiality of film	97
6	Screen space II: factual television	139
7	Screen space III: depiction and the space of interactive media	156
8	Conclusions	198
	Bibliography	207

1 Introduction

1 About the thesis

The thesis comprises the following chapters:

1 Introduction
2 Approaches to the study of space
3 Depth perception and depiction
4 Spatiality and realism
5 Screen space I: the spatiality of film
6 Screen space II: factual television
7 Screen space III: depiction and the space of interactive media
8 Conclusions

'Media space' in this thesis refers to spaces which may be *seen* in the planar media of still pictures, film, factual television, and interactive digital media presented on computer displays. It inevitably touches on configurational approaches to such space, but gives priority to the subject of depiction. While substantial discussion is given to the meanings which viewers or users take from planar pictures, it is not concerned with those spaces which are essentially metaphoric, such as 'hypertext', 'information space' or 'cyberspace' (though of course these often also have visual representations).

In 1992 Gaver pointed out that as technologies such as video, audio, or computer change, so do the kinds of space that it is possible or desirable to make. He used the term 'media space' and suggested that its analysis would highlight possibilities for design (**Gaver 1992**). This is also the rationale for the present thesis – to stimulate fresh approaches to spatial design and to the spatial aspects of depiction – even though the interpretation of media space here is, as indicated, different from the one Gaver had in mind.[1]

The thesis is that picture-making (to be defined) is best conceived of as a pragmatic activity in which the objectives intended for the artefact have equal weight with (and often take precedence over) the relation between the picture and what it depicts. It is concerned with the visible formal aspects of pictures, rather than with, for example, metaphors of space.

The method used is to analyse how space is depicted and used in figurative representations. Analytical description is informed by texts from a number of normally discrete literatures, since my subject is not (yet) a defined discipline.

To bring out the pragmatism of picture-making offers two important advantages. It makes it easier to explain the multiplicity of kinds of pictures which have been made and continue to be made. Snyder (**1980 p246**) suggests that 'the grab bag category of realistic pictures will forever defy general analysis' but this is just what I aim to do. It also, I hope, offers the designer a way of thinking about the demands of unfamiliar

[1] Gaver's paper related to network-linked work locations supporting a variety of media such as video, audio and digital data.

media, encouraging the development of new pictorial solutions. I emphasise in the final chapter the idea that interactivity, like any other change of use of pictures, permits – and requires – new spatialities.

I ask the reader to tolerate terms which have only cursory initial definitions, on the understanding that their full definition is part of the subject of the thesis itself. For example, to say that the working definition of a picture is that it is 'a planar representation of a scene' raises the questions of what a *scene* is and how depiction *represents* it. These questions are an important part of this enquiry.

The motivation for the research is partly practical, arising from my work in teaching designers of digital interactive media. In considering various computer-based products and projects, I have found myself thinking that, whereas pictures, film and television seem to make expressive use of space, both as subject and medium, interactive media, certainly so far as they deal in pictures, generally do not. This has led me to enquire what this apparent expressivity is and what its sources are.

2 The chapters

Chapter 2: Approaches to the study of space

This chapter offers a brief account of two approaches to explaining how depiction works, which in some respects are opposed. One gives predominance to acultural factors while the other claims a high level of cultural conditioning in the way that pictures are made and perceived.

I note perceptual theories of how spatial configurations are interpreted, particularly concerning the extent to which this is acultural, and show examples which seem to suggest a significant role for cultural interpretation even of simple artefacts.

This is one of the few times in the thesis when I touch on non-pictorial spatial configuration (the others are in connection with television and digital interactive media), which is outside my scope. The study of such configurations, whether through perception science, theories of diagramming or of design more broadly considered, is a huge subject in its own right.

The other approach noted is that of theories of visual culture. This is included to establish a context for the consideration of space and specifically for a discussion of relativist views of the geometry of depiction in Chapter 4, and to introduce three ideas which I make use of throughout the thesis.

This chapter contributes only a little which is original to the research, but forms a necessary foundation for what follows.

Chapter 3: Depth perception and depiction

This chapter concentrates on the use of pictorial images to suggest depth. Its main purpose is to decompose the single concept of depth depiction into a range of

techniques, exploring how each offers different affordances (defined below) through the resulting image.

In defining pictures, I propose a simple framework which characterises the components of picture-making as each in its own way representational. In particular I distinguish decisions made about the *model* – that which is to be depicted – from those concerning the *view* and its representation through *picturing*. This framework is used and refined throughout the thesis.

One of the most obvious ways in which pictures and the world they depict differ is that pictures are generally flat, while the world is not. This chapter takes the thirteen depth cues of visual perception proposed by JJ Gibson (**1950**) and applies them to depiction. This may seem an odd approach given that Gibson consciously formulated his cues in opposition to an established tendency to consider vision in pictorial terms. However, by highlighting the differences between perception and depiction – and by anatomising the general concept of depth – it is beneficial to the thesis presented here, especially because it emphasises those aspects of depth depiction which are not to do with geometry and suggests that selectivity in the use of pictorial depth cues may have advantages, which I begin to evaluate.

In analysing pictures in relation to their evocation of depth a number of questions are raised which are pursued through the remainder of the thesis.

The application of Gibsonian theory to pictures is an original approach which helps prepare for the following chapter on realism, not only by emphasising historic and contemporary selectivity in depth depiction, but also by hinting at the difficulties of saying quite what it is that pictorial media are supposed to imitate.

Chapter 4: Spatiality and realism

This chapter is pivotal in the thesis, following the analysis of static pictures and preceding that of dynamic screen-based media. I ask what the criteria are for considering a picture to be realistic. Contrary to the advice of Gombrich (**1973**), I argue that the illusion that a picture is an actual scene (*pictorial illusion*) offers the best means of testing claims of pictorial realism.

Analysing both texts and artefacts, I critique the generalised notion of realism and suggest replacing it with an approach based on multiple realisms. These different realisms tend to be promoted by different kinds of depth cue and have varying results for the user of the artefact (I distinguish particularly between the informational and the affective). I argue that these multiple realisms cannot usefully be subsumed into one 'super-realism' which approaches the illusion of looking at actual scenes, principally because realisms are selected to serve particular objectives, even when the picture-maker believes that some unitary notion of realism is employed.

I evaluate the claim of PI-realism (defined as that realism which could fool the observer of a picture into believing that they see a real scene) to present a true spatial geometry of scenes. I show that the idea that many projection systems have equally valid claims to be accounted realistic is, on the criteria of pictorial illusion, incorrect. In the process I highlight errors made by some major theorists. However, I proceed to question the *usefulness* of this correct projection when tested against various demands. These include difficulties of implementation, failure to correspond to more subjective aspects of vision, and mismatch to the many functions which pictures are intended to perform.

I show that attempts to accommodate aspects of time in still pictures are an important influence on their form, and this leads me to propose an alternative conceptualisation of realism, *visual experience* realism, which in a variety of ways aims to make looking at a picture evoke aspects of looking at scenes (not the inert visual stimulus of the scene itself). This alternative realism I argue is both more limited (since it does not unequivocally match an external referent and is qualified by cultural codes) and more expressive (since it allows the design of pictures to be attuned pragmatically to their objectives).

The argument thus runs that though there are projection systems which have a particular claim to be regarded as superior in matching the optical stimulus before the eye, for most picture-making this is irrelevant because of the failure of such images to capture subjective and time-based aspects of vision (*failure of correspondence*) and the likely mismatch of such pictures to the objectives which they are intended to fulfil (*mismatch of function*).

I argue that the 'limitations' of still pictures have been turned to strengths and that this is a mark of a *mature* medium. This theme is pursued through the remaining chapters.

The framework proposed in the previous chapter which links the *objectives* of a picture to an *idea*, instantiated in a *model* which is *viewed* and *pictured* is now reconsidered to take account of the way in which for a picture, and, it will be seen, for film, every aspect can be considered in some sense pictorial, since the model and view may be, usually are, configured with a view to achieving a certain pictorial outcome.

Chapter 5: Screen space I: the spatiality of film

Film is unlike natural vision in being framed, as still pictures are, but it does offer motion and other time-based phenomena, so remedying one of the key 'failings' of pictures. This might mean that film is much more like natural vision. However, I demonstrate that the broadly mimetic intra-frame qualities of film do not extend to its inter-frame spatiality. This forms another stage in the argument concerning pragmatism. By analysing intra-frame space (the variables of viewing and picturing) and the larger spatial practices of film, I demonstrate that the key requirements on

film are concerned with authorially determined showing. This is subject to many demands, of which any matching to natural vision is relatively unimportant.

However, the spatiality of the mainstream fiction film is not in general designed to be *perceived* as authorially contrived, so that the suppression of overtly artificial spatial practices in mainstream film is one of its distinctive traits. I use three examples – Welles' *Citizen Kane*, Gance's *Napoleon* and the films of Peter Greenaway – to draw out the spatial differences between those kinds of film which belong to the fictional mainstream and those which do not.

The objectives-to-picture framework is applied to film. I show how, as with painting, backtracking is an essential aspect of the process: though as a pipeline of physical process it is clear that the set is built, then populated with actors and finally viewed and pictured, every aspect is designed in the light of the final image (and normally in the context of the adjoining shots). In that sense I argue that fiction film could be considered wholly pictorial.

The most important function of this chapter is to demonstrate how the spatial practices of film are bound up with its being an authored narrative. This relationship of authorial control to the spatial practices of film hints at the difficulties explored later of transferring those practices of film to media which are interactive.

I argue that film should be regarded as a mature genre with a high level of expressivity – it is a genre in which any spatial practices which do not suit its objectives have been expunged. This is not to suggest that its characteristic spatiality will not continue to develop, since it is part of my argument concerning maturity that technological change and audience expectation – even of the supposedly 'natural' – develop together.

This chapter offers a new analysis of film in spatial terms and in so doing reveals aspects of the medium which have not previously been highlighted, probably as a result of the general preoccupation with film's temporal rather than spatial characteristics.

Chapter 6: Screen space II: factual television

The spatiality of television has hardly been studied; this chapter is novel. I trace the recent history of those spatial practices ousted from film-making which continue to flourish in factual television, and consider why this has occurred. Of all the many genres of television, factual programming is selected here because it most noticeably uses those spatial practices rejected in the fiction film. This reinforces the argument that spatial practices are closely allied to genres, which are in turn bound up with the objectives of the artefact.

Some of the characteristics of factual television are inevitable given the nature of its raw material which usually does not offer many opportunities for the construction

and selection of model and view to achieve a particular pictorial outcome. However I argue that more important than this is the key difference between the spaces of fiction film and factual television – that the former aspires generally to seem like natural vision (however artificial it may really be) while factual television is content to be perceived as explicitly 'showing'. Nevertheless, even within this one genre, variant spatial practices can be discerned, which once again reflect the different objectives of various programmes.

I claim that the spatial practices of factual television, while opposed to those of the fiction film, are just as well attuned to the objectives of each artefact, and that this marks out factual television also as a mature medium.

Chapter 7: Screen space III: depiction and the space of digital interactive media
In this penultimate chapter I apply the approach developed so far to the various spatial practices of digital interactive media, which have hitherto not been studied. I propose a taxonomy of seven types reflecting the range of current practice. Based on the criteria developed in the thesis, I describe this practice as by-and-large 'immature', though I offer an explanation as to why this is inevitably the case. Within the taxonomy, I discern the two broad tendencies already outlined – namely that some artefacts aspire to seem like simply seeing, while others are overtly presentational. As also for film and television, I emphasise that spatial practices belong not to media or technologies but to genres, and that one should therefore expect that there will be increasingly divergent spatial practice as the genres of digital interactive media evolve.

Unlike the previous chapters, this one concludes with suggestions as to what may, and perhaps should, happen in the development of the spatiality of this medium, based on the arguments of the thesis as a whole.

Throughout, I take pains to avoid neat theories which fail to fit the evidence of artefacts. The arguments are therefore couched in terms of tendencies and characteristic qualities rather than rigid prescription. The 'messiness' of spatial practice, arising from its pragmatism, is highlighted on many occasions. Nevertheless, I am confident not only that the main argument is clearly made but that other important advances are made in the development of a specifically spatial approach to depiction.

3 Concepts and definitions used

A small number of terms which recur throughout the thesis are defined and briefly discussed here.

Pictures
The thesis is largely concerned with the nature of depiction, so it would be premature to define the term *picture* here. However it is probably useful to make clear at this stage that I do not differentiate between representations which are derived from

scenes in the world really observed by the picture-maker, those derived from the imagination and those which are some combination of the two. I believe it will become clear in the course of the thesis that the formal issues concerning the relationship of the depiction to the depicted are the same whether the depicted matter is fictional or factual, so no purpose would be served by making this distinction.

Automorphism

The term *automorphism* arises in discussing pictures and realism. It refers to matching aspects of a representation to corresponding aspects of whatever is represented. In an automorphic representation, colour of ink might be used to represent the colour of the thing it represents. There are degrees of automorphism, so the colour of the ink might match the hue of the object, but not its tone. Spatially, in an automorphic representation a symbol being to the right of another might stand for the thing it represents being to the right of the other – as is normally the case in pictures (**Currie 1995 p97**). If a picture could be a perfect copy of a scene then it would be completely automorphic.

Affordance

Representations, I argue, are made for purposes, not for their own sake: the artefact lends itself to particular ways of thinking and makes possible certain operations. As has been said of the spatial organisation of text, 'one can operate on these representations' (**Goody 1987, p187**) and in the case of interactive systems this is literally true.

The concept of affordances arose in Gibson's work on perception (**1979**) and has been popularised by Norman (**1988**) in relation to design. It refers to the way in which any object has preferred 'readings' – inferences the user makes as to how the object can be employed. So, for example in looking at a simple tool like a spade we are almost involuntarily drawn to the idea that one part is for hands to hold, while another, the blade, is for piercing and cutting. The parts of the artefact *afford* particular operations.[2]

Norman has complained (**1999**) that affordance has become a mindless talisman in the HCI literature. I use it in this thesis because it emphasises that what the designer does is to make some readings *more likely*, not certain (the user might misinterpret the form of any design, however simple). It implies the inevitability of some error in interpretation (a key point in **Norman 1988**), emphasising that the designer is not *making meaning*, only favouring the apprehension of certain meanings among others.

Diegetic space

Rimmon-Kenan (**1983**) notes how the distinction between Plato and Aristotle's *mimesis* (often translated as *showing*) and *diegesis* (*telling*) had come by the end of the nineteenth century to mean, on the one hand, the direct presentation of events and conversations in which the narrator seems to disappear (as in drama), and, on the other, presentation mediated by a narrator who, instead of directly and dramatically

[2] Strictly speaking there are two different aspects to affordance: (1) what can be done with an object, (2) *what its appearance suggests* can be done with it. The latter is the important one here.

exhibiting events and conversations, talks about them, sums them up, and so forth. Since then, the term *diegesis* has been widely adopted in film theory with a rather different sense: it now usually means the space or world which is accessible to the characters in a film, so that for example Giejgo says (**2001**): 'Events and knowledge known to the characters within the plot and story are diegetic. Therefore characters only perceive diegetic material. Audience[s] can perceive everything a text has to offer – including non-diegetic material.' While Armes complains that this use of the term is 'misleading' (**Armes 1994 p12**) he acknowledges that it is now conventional and I have adopted the newer meaning here. For any kind of depiction, it will prove useful to be able to distinguish between *diegetic* space – that which is occupied by the world depicted in the image – and *non-diegetic* or *extra-diegetic* space which is occupied for example by a set of controls for interacting with the world. In the case of a painting, the artist's signature is usually in the non-diegetic space of the canvas – but occasionally is made diegetic by being incorporated into the scene itself.

Function and affect in design

Certain assumptions are made in this thesis about the nature of design, particularly that both function and affect are important in most kinds of designing. If we consider various kinds of artefact, it is clear that though many have a functional purpose (or several purposes) their form is not limited to functional considerations. Clothes must normally keep the wearer warm and dry, yet this tells us almost nothing about the innumerable different forms of garment which are worn. Legibility in the design of a newspaper is presumably essential, but the reason why one newspaper looks different from another is not to do with legibility, but with the associative qualities of the type-face, layout, proportion of picture to text, and so forth. These are affective aspects of a design, intended to do something to the user, to alter the relationship between the user and the artefact.

It is therefore useful to think of the objectives of a design not just in functional terms but as including the affective aspects. For pictures, these may be crucial in determining the spatial attributes. For example, to see an object depicted stereoscopically may not provide any significant *information* which is not visible in a monocular view, but it will alter the *affect*: users may have a stronger sense that the object exists in the same space as themselves. Such a case demonstrates that information and affect cannot in fact be rigidly separated. For example, if the sense of immersive presence in a virtual environment enables the user to achieve a better understanding of what some corresponding real experience is like, it could be argued that the user is thus more fully informed. Film space helps demonstrate that informational and affective aspects can become inextricably bound up in a single spatial device, so that, for example, moving the focal plane (pulling focus) is informational since it allows something else in the scene to be more clearly observed, but it has a strong effect on the viewer-subject relationship. The difference between function (in this case the provision of substantive information) and affect (an altered relation to the representation) is thus one of emphasis rather than a categorical distinction.

Expressivity

The word *expressivity* is often used artistically to suggest some rather vague idea about how a picture, or a film, or a digital artefact, expresses its meaning. On the other hand, in the context of information design, Mackinlay (**1986 p114-6**) proposes a strict test of expressivity that representations must 'encode all the facts in the set and encode only the facts in the set'. This is a view which I suspect is simplistic even in its original context,[3] but which certainly does not extend well to other kinds of representation. I show in the next chapter that all but the most trivial visualisations can only be understood on the basis of prior knowledge and that the observer employs that prior knowledge to decide *which* spatial attributes of the representation are meaningful. It is not possible to say categorically: these are the facts and a design conveys these and no others. Such a view also leaves out of account the affective aspects just discussed. Since a design includes the relationship between the user and the artefact it cannot be conceived as purely the transmission of facts. A more subtle definition of expressivity is developed in this thesis.

[3] Though in some respects Mackinlay's test of expressivity seems reasonable, it implies that all the structures which could legitimately be incorporated in a representation inhere *in the data*. While it is reasonable to say, as Tufte (**1983 p55-77**) and Wainer (**1997 p22-25**) also do, that arithmetically significant spatialisation should not be used in the absence of data (so that three-dimensional visual representations are demonstrably misleading when only two dimensions of data are available), Mackinlay's dictum implies that all the facts are pre-configurational, that all that can be found is in the data itself, independent of context. I suggest that the context shared by designer and user, which is essential to the user's understanding, calls the purity of his view into question. Mackinlay takes a view of the designer as one who *transmits* pre-existent facts, rather than as an agent who interprets and discovers.

2 Approaches to the study of space

1 Introduction

Taking the concept of space in general terms, a number of questions arise: Is space perceived the same in all times and places? Is there an idea of space which is universal? Does the conception and depiction of space relate to the values of the culture that produces it? I summarise in this chapter contrasting approaches to space which help create a context for the following chapters.

Looking for constants in our relationship to space, Lakoff and Johnson highlight the preponderance of spatial metaphors in language (**1980 p14-21**) and suggest that since our world is spatial, we are contained in space and are ourselves containers of it, the notions of space which pervade our everyday experience are likely to be shared across cultures. Similarly Harrison and Dourish emphasise those aspects of the spatiality of the world which seem to offer a universal frame of reference. They suggest that 'dealing with physical structure, topology, orientation and connectedness, *spaces* offer opportunities and constraints. *Places*, on the other hand, reflect cultural and social understandings,' (**Harrison and Dourish 1996** *emphasis added*). In contrast, Hall (**1959**, **1966**), arguing that space is fundamental to non-linguistic communication, identified culturally determined variations as well as constants. Recent studies of the spatial dynamics of social organisation again emphasise the interrelatedness of spatial and social factors (**Dunbar 1996** *passim*). Markus, studying the effects of principles of social organisation on the design of buildings has suggested that 'there is no a-spatial society and no a-social space' (**Markus 1993 p13**). He documents the influence of the Panopticon, also cited by Foucault (**1970** (**1966**)) as an emblem of the 'politics' of seeing which has inspired many other writings within the ambit of *visual culture* (see below).

I shall argue that the form of pictorial representations has a strong relation to their purposes, which presumably cannot fail to reflect the values – and not just the graphic capabilities – of the culture within which they are made and used. Spaces, built or depicted, 'reflect cultural and social understandings' just as places do. It is however also useful to consider the invariant aspects of spatial awareness.

2 Components of spatial perception

Perception science has a need to identify how we makes sense of what we see. A principle of all approaches to perception is the assumption that the retinal image is *in itself* of no use: only some sort of comprehension of the image can provide advantage.[1] At what point this comprehension occurs, and under what influences, is a matter of dispute. Nevertheless most models incorporate aspects which are data-driven, determined by inherent mechanisms of the visual system, and aspects which are hypothesis-driven and are subject to influence by prior knowledge (Figure 2.01). Gregory (**1987**) dates the assumption of a hypothesis-driven component to the work

[1] This comprehension may not amount to much, yet still serve its purpose. Marr (**1982 p32-4**) cites the limited spatial 'comprehension' of the housefly, which is nevertheless sufficient to trigger appropriate actions. Vision thus extracts from images of the world information which is useful and relevant (***op cit* p31**).

of Helmholtz (1821-94). Even Marr, who emphasises what can be achieved by the visual system without recourse to hypotheses, proposes a three-stage process model (**1980**) in which, following the generation of a *primal sketch* and a *2½-D sketch*, the recognition of a *3-D model representation* is influenced by experience of typical object shapes.

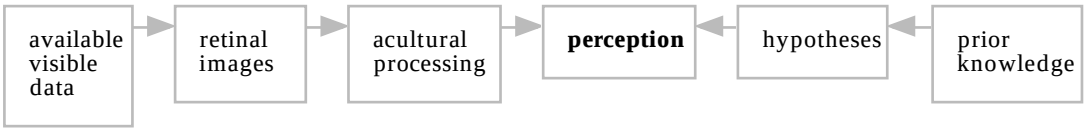


Figure 2.01. Conceptual relationship of data-driven and hypothesis-driven aspects of perception

The sense-making work of the visual system which is not influenced by prior knowledge has been described in terms of *gestalt*, a term coined by Wertheimer (1880-1943) to capture the idea that some sort of pattern is derived from sense-data before any cultural understanding is applied. Whether or not framed in terms of gestalt theory, there seems to be a broad measure of agreement that some of the visual system's organisational abilities are so fundamental that they are not significantly different across cultures (**Marr 1982**, **Hochberg 1980, Gregory 1987, Barlow 1990**, **Bruce, Green and Georgeson 1996**). Later in the thesis, the difference between two kinds of image – the real image projected by light on the retina and the 'virtual image', an internal representation which the visual system as a whole creates from the retinal image – will create many difficulties in establishing the meaning of visual realism.

Lansdown, in a summary of gestalt perceptual theory (**Boyd Davis, Lansdown and Huxor 1997**), highlights *smooth continuation*, *proximity*, *similarity*, *orientation*, *closure*, *relative size* and *common fate*, which I briefly define here. Most are accessible in static stimuli, while one – common fate – arises only in relation to movement. One can speculate on ecological origins for these phenomena, an approach to perception promoted in particular by Gibson (**1979**) which has informed subsequent thinking on perception.

In *smooth continuation* the observer groups together in a single structure those parts which seem to align or continue smoothly. This may be part of the process by which we segment the retinal image into objects: parts which align are seen as belonging to a single object. Solso suggests that it is also related to ecologically valuable trajectory prediction, since good continuation in object paths tends to fit well with the laws of the physical universe such as gravity and the conservation of momentum (**Solso 1994 p95**).

In the case of *proximity* the observer groups those parts that are closest together. This presumably assists in identifying both single entities which have modular parts and groups of objects which are likely to behave in a coherent fashion, such as flocks and other groups. In a set of objects, *similar* objects will be perceived as belonging together. Under suitable circumstances, this will cause the pattern of the objects to stand out as an identifiable form. Objects may be perceived as parts of a group by virtue of their distinctive *orientation*.

Closure 'completes' intermittent parts of an enclosing shape. As Solso indicates, enclosed shapes seem to have greater concreteness than unjoined marks (**Solso 1994 p96**). This touches on *figure-ground* phenomena: given a small shape within a large one, the observer tends to see the smaller as a figure against the larger background rather than as a pierced hole in the larger shape. When there is little difference in the size of the parts, ambiguity can result as in the familiar faces / vase illusion.

The tendency to see stimulus components as belonging together, as an entity or as a group, is accentuated when movement is introduced: items that move together, have a *common fate***,** are seen as grouped together. Though a shape may not be discernible in a static pattern of dots, when the dots move in a coherent way, the shape becomes apparent.

Limitations of the scope of gestalt principles

It is easy to observe difficulties where a visual configuration has been made ignoring gestalt principles. In Figure 2.02, study reveals that there are two separate '*and*' *gates* in the top row of the blue diagram, but this is not the user's initial impression because proximity causes grouping of the top row as a whole.

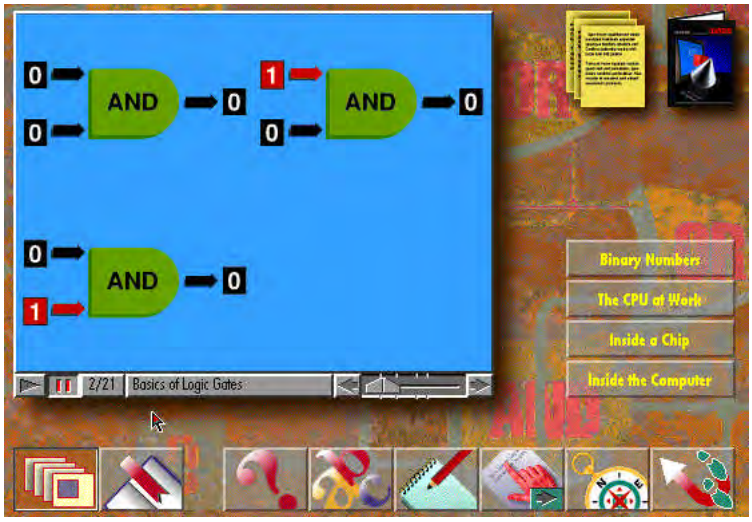


Figure 2.02. A gestalt principle ignored

'How computers work' CD-ROM (Time-Life Books with Warner New Media , 1993) (detail).

When the designer ignores basic principles of spatial organisation, the user has difficulty making sense of what is shown.

In some cases one gestalt principle will conflict with another. For example, Gregory shows a case[2] (Figure 2.03) where the tendency to organise dots into rows through smooth continuation overrides the grouping effect of proximity (**Gregory 1970 p20**).

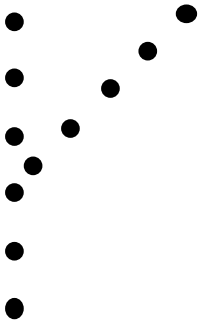


Figure 2.03. Competing organisational principles.

From Gregory 1970 p20.

The lowest dot of the diagonal is close to two dots of the upright. However it is the two rows which are perceived as primary structures, not the cluster of three dots.

While it might be argued that the benefits of acultural perception are still available in

[2] Gregory does not describe it in gestalt terms.

a case such as this, since presumably *all* observers will prioritise continuity over proximity, the suspicion arises that no safe prediction could be made for the indefinite number of variant designs where these two (or other) gestalt principles might compete. Marshall and Shipman (**1995**), in developing designs for a visual interface to hypertext in which users could move objects about in order to express relations between them, found not surprisingly that users expressed these relationships by using proximity, alignment and similarity in combination.

Cultural conditioning tends to invade the interpretation even of very simple visual structures. For example, two small diagrams are used together on the door controls of certain UK trains (Figure 2.04). The upper symbol denotes opening, while the lower denotes closing, of the double doors. It relies on the learned recognition of the angled marks as arrow-heads (denuded of their shafts) which are to be understood as indicating the direction of travel of the doors which will result. No doubt most users are familiar enough with this graphic convention to recognise their meaning correctly. However, considered purely in gestalt terms, the upper symbol has a unity which might be analogous to the closure of the doors, while the lower symbol conveys a contrasting impression of fracture. Only the user who has learned the diagrammatic conventions of arrow-heads is likely to make a correct interpretation, even of these extremely simple spatial configurations.

Figure 2.04.
Graphics on buttons for operating train doors
Only the user familiar with the diagrammatic conventions of arrow-heads is likely to make a correct interpretation of these configurations.

We may underestimate the extent to which learned knowledge is required. At the beginning of modern data visualisation,[3] in the 'lineal arithmetic' of Playfair (1759-1823) (**Tufte 1983 p32, Bertin 1973 p222**) a method is described which is so simple that it seems to rival the acultural qualities of gestalt principles: larger marks stand for larger quantities – 'presenting to the eye a figure, the proportions of which correspond with the amount of the sums intended to be expressed' (Playfair 1801 p*ix-xi* quoted in **Twyman 1986 p223**). However when Playfair continues his explanation it becomes clear that this is not so:

> Suppose the money received by a man in trade were all in guineas, and that every evening he made a simple pile of all the guineas received during the day, each pile would represent a day, and its heights would be proportioned to the receipts of that day; so that by this plain operation, time, proportion and amount, would all be physically combined.
> Lineal arithmetic then, it may be averred, is nothing more than those piles of guineas represented on paper, and on a small scale, in which an inch (suppose) represents the thickness of five millions of guineas, as in geography it does the breadth of a river, or any other extent of country.
>
> Playfair 1801 p*ix-xi* quoted in **Twyman 1986 p223**

[3] Playfair introduced the line graph, bar chart and area chart (**Twyman 1986 p220**). His work was not entirely unique: Johann Heinrich Lambert in Germany also used line graphs in his scientific writings from the 1760s, both for analysis and display (**Twyman 1986 p221, Tufte 1983 p32**). Tufte nominates a seventeenth century example as the first statistical graphic (**Tufte 1997 p15**).

As described, the method owes a lot to depiction (the aspect of geographic mapping to which Playfair compares it being also essentially a pictorial rather than a symbolic one), and it is not impossible to imagine that an observer wholly unused to graphic visualisation might grasp that the heights of marks on the paper correspond to the quantities being measured. However no such assumption can be made about the horizontal axis, for all that Playfair implies that the piles of guineas would be set out in chronological order. The idea that time can be represented quantitatively, using the same graphic variable as actual quantities, is entirely cultural: it does not operate at the basic, irresistible level of gestalt principles (hence perhaps the necessity of Playfair's 'manifesto' and explanation). This relation between time and quantity used by Playfair is perhaps best considered a metaphorical one. Though Lakoff and Johnson have argued (**1980 p p14-21**) that some of the commonest metaphors, including the spatialisation of time, arise from our shared physical experience of the world (so they might not be cultural), their examples are all spatialisations of time *in relation to the observer* (such as the use of the concept 'look forward' to mean both looking to a later time and looking in front of oneself) rather than a representation of time such as Playfair's dimension on paper, which has its own spatiality independent of that of the observer.[4]

Playfair's contemporary Joseph Priestley (1733-1804) used proportional graphic representation of time as the basis of his work (Figure 2.05) and this may have been a direct influence on Playfair (**Twyman 1986 p216**).[5]

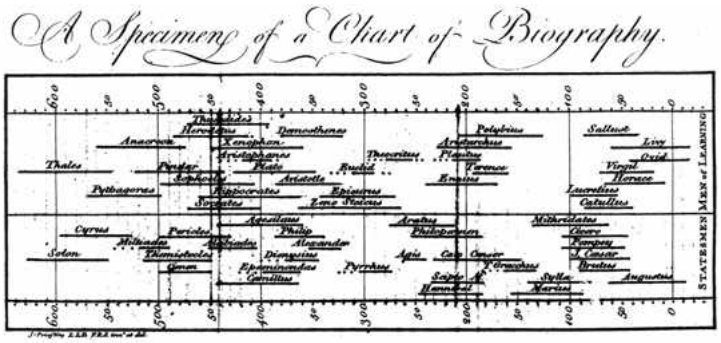


Figure 2.05.
Part of Priestley's explanatory leaflet for his timeline,1765, reproduced in **Twyman 1986**

While gestalt principles clearly facilitate the reading of such timelines (such as the noticing of clusters at times when several notables lived, compared with fallow periods), many of the visual properties must be 'decoded' to be understood. There is very little of value in this representation which can simply be 'seen' in the sense, say, that one can see planar relationships in a geographical map. The observer must employ prior knowledge to decide *which* basic spatial attributes of the representation are meaningful. Visual clustering inevitably takes place whether the grouping is horizontal (meaningful) or vertical (meaningless) since the perception of clustering cannot differentiate in terms of axes.[6] Likewise while users should infer meaning from the vertical *alignment* of two items, they should not infer anything from horizontal

[4] In mapping other kinds of data, acculturated metaphor is even more evident: for example in the ideas of mapping a 'high' temperature, or a 'high' note in music.

[5] His innovation in graphic visualisation was the development of the timeline (**Twyman 1986 p216**). He also published important research into the nature of electricity and contributed to the discovery of oxygen.

[6] If it makes any sense to consider the perception of clustering in coordinate terms, it could be considered as multi-polar so that the nearness of all other objects to each object is measured.

alignments. Even though users need not necessarily understand the intentions of the individual who devised this specific representation, they must be in a position to use knowledge of the communicative intentions normally embodied in this kind of diagram.

My aim here is not to criticise the quality of this particular information design, but to indicate the inevitability of invasion of the semantic field by unwanted meanings. Even quite simple graphic organisations such as a timeline cease to be able to rely on the foundational predictability associated with gestalt principles. In the somewhat unpredictable relationship between culture and perception, prior knowledge is necessary at a surprisingly low level of interpretation.

3 Visual Culture

An approach grounded in cultural history offers a relativist view in strong contrast to the universalising aims of perception science. While the defining characteristic of theories of *visual culture*[7] is a belief that the visual artefacts which a society produces and uses indicate and inform its values and beliefs, it has also come to stand for a number of related theories, including that visual realism is a social construct and that the shared attitudes of a culture are 'invisible' to those who hold them. In this thesis I make use of both ideas, though I take issue with the extreme view that every aspect of figuration is culturally determined and that therefore no pictorial geometry has a better claim to be considered correct than another (Chapter Four). I often use the idea that the familiarity of particular forms of representation – what Barthes (**1977 p17** *passim*) would call 'codes' – leads to their becoming invisible or transparent. I also borrow the idea, implicit in most writing on visual culture, that intention can be embodied in an artefact even when no conscious intending, and no particular intending individual, is implied.

Relativism of spatial concepts

There are many references in this literature to alleged differences in perception, conceptualisation and estimation of space. Differentiation by culture is generally seen in qualitative terms, while quantitative differences in the estimation of space are usually linked to era.

Perhaps the earliest visual culture text predates the term itself: Panofsky's *Perspective as Symbolic Form* (**1991** (**1925)**). This essay set the agenda in two ways: as the title suggests it explored the relationship between planar representations of space and the cultures which make them; it also contains a significant error concerning perspectival depiction which has since been repeated by others and with which I deal in Chapter Four.

Harvey considers Durkheim to have been the first anthropologist to suggest that different societies produce fundamentally different conceptions of space and time

[7] Alpers is credited with the dissemination of the term in 1972 (**Evans and Hall 1999 p5**) though Alpers herself attributes its origination to Baxandall (**Alpers 1983 p*xxv***).

(**Harvey 1996 p210**). In the linguistic field, the suggestion that they might be culturally determined arises in the work of Sapir (1884-1939) and Whorf (1897-1941). The Sapir-Whorf hypothesis is now cited as fact in some sources (for example **Kress and van Leeuwen 1996**). A favourable view of Whorf's work is also found in a paper by Hopgood (**1993**), who was struck, while developing standards for specifying time-based multimedia, by the widespread tendency to make categorical assertions about both time and space. His conclusion (***op cit* p3**) was that 'Europeans have a notion of time and space that is generally assumed by them to be universal. This gratuitous assumption is naive, arrogant and wrong.' As Hopgood suggests, this implies shifting the requirements of visual representation, even in figurative pictures, from a simple (if problematic) attempt to capture the truth to a more functionally directed agenda.

The key argument of the Sapir-Whorf hypothesis is that *perception* of space is determined by culture, and particularly by language. Pinker aligns himself strongly against this nurture-based view: 'No matter how influential language might be, it would seem preposterous to a physiologist that it could reach down into the retina and rewire the ganglion cells' (**Pinker 1994 p62**). This is surely a false antithesis: it is not necessary for culture to alter physiology in order to interfere with perception, unless visual perception is considered synonymous with vision, in other words as having no hypothesis-driven part but being entirely data-driven. However he does usefully point out the circularity in Whorf's arguments, in that language is the only evidence Whorf has to demonstrate that people perceive differently!

Quantitative differences in the *estimation* of space are reflected in the tendency widely remarked in industrial societies to see space as becoming compressed as a result of the ease with which it is traversed. Thrift (**1996 p264-265**) considers this compression to have been a phenomenon of stagecoaches, railways, bicycles, the post and the telegraph: the railway age has many such references.[8] For Heine (quoted in **Schivelbusch, 1978 p34**) 'the elementary concepts of time and space have begun to vacillate. Space is killed by the railways. I feel as if the mountains and forests of all countries were advancing on Paris.' The Quarterly Review of 1839 predicted that 'As distances were thus annihilated, the surface of the country would, as it were, shrivel in size until it became not much bigger than one immense city.' (**Schivelbusch, 1978 p32**). For Williams (**1852 p284-85**) 'the extremities of the island are now, to all intents and purposes, as near the metropolis as Sussex or Buckinghamshire were two

[8] Unfortunately Thrift like a number of writers in this area, builds fantastic claims on the basis of these reasonable observations, *viz:* 'even if in some mystical past it had been possible to analytically separate space and time, in the contemporary world the notions of space as enclosure and time as duration are unsettled and redesigned as a field of infinitely experimental configurations of space-time' (**Thrift 1996 p285**). In recent years the idea of spatial compression through technology has become a feature of *cyberspace*, a term generally denoting metaphorical spaces (Gibson, coining the term in 1986, gave a kind of visual equivalence to the metaphor as 'bright lattices of logic unfolding across the colourless void', **Gibson 1986 p10**). According to Taylor and Saarinen (whose experience of cyberspace at the time seems to have been confined to the use of email) 'omnipresence descends from the heavens and becomes actual on earth' (**Taylor and Saarinen 1994 p2**). There are striking similarities between the polemic of cyberspace and that of the railway age. As Marvin points out (**1988, p193**), in the nineteenth century it was argued that, 'instantaneous electric communication augured a universal language, usually thought to be English, and global harmony [...] the devices which social imagination constructed and then reacted to sometimes actually existed, but just as often were entirely imaginary'. New technologies were seen not only as annihilating space and time but also difference. For Ascott (**1994**) 'cyberception not only implies a new body and a new consciousness but a redefinition of how we might live together in the interspace between the virtual and the real.' Like Marvin, Standage notes the hyperbole of spatial compression in relation to the telegraph (**Standage 1998 p125**) and the similarity of the benefits claimed for the telegraph and for cyberspace.

centuries ago.' Massey (**1993 p60-61**) has pointed out that this notion of time-space compression needs differentiating socially: it is hardly reasonable to suggest that the sensation of spatial compression was the same for all at a given historical period, regardless of wealth, gender or other factors. This returns the discussion to qualitative differences of spatial interpretation.

Such relativistic thinking has directly influenced some kinds of visual representation of space. For example, whereas geographers traditionally saw themselves as dealing in the appraisal and mensuration of the physical landscape, they have taken an increasing interest in the semantics of space and this has affected the kinds of maps which are made. According to Jackson (**1993 p207**), geography has been 'remodelled, remade and rethought' in a shift away from the 'obsession with landscape'. The concept of 'relative space' has dominated feminist and other challenges to traditional 'absolute' spatial geographies (**Rose et al. 1997** *passim*). Cartesian distances are replaced by models representing, for example, accessibility by particular users. In fact the term 'spatiality' has, following a suggestion by Soja in 1985, been used by geographers mainly to denote this kind of socially produced and interpreted space (**Rose et al. 1997 p4-7**).

In some respects this has led to an attack on map-making in general, leading Pickles (**1992 p193**) to complain that 'the theory of maps has received comparatively little attention amidst the burgeoning literature of the new theoretically informed geography'. One innovative visual form which seems however to have been accepted into the mainstream is the 'cartogram' used by Kidron and Segal in 1981 (**Dorling and Fairbairn 1997 p146-150**) in which countries of the world are assigned approximately their conventional Cartesian locations but are scaled according to other criteria than topography, for example according to per capita CO_2 emissions (**Smith 1999 p98-99**).

Of course 'traditional' geography has its own history. For Harvey (**1996 p239**), the introduction of the Ptolemaic map into Florence in 1400 is fundamentally linked to the emergent needs of trade and commerce, of property and territorial rights which differ from those of the feudal world. However, the modernity (or postmodernity) of such concerns is often overstated: it is difficult to believe that anyone has *ever* held the 'traditional theory of maps as unproblematic mirrors of nature' which Pickles suggests (**Pickles 1992 p193**).

The 'hegemony of vision'

A particular characteristic of much visual culture theory is its tendency to attack *any* notion of visual realism, offering in its place an entirely relativistic view (I return to this in Chapter 4). There is a tendency to regard the alleged 'hegemony of vision' as an actual conspiracy. Thus Crary states that 'The standardisation of visual imagery in the nineteenth century must be seen then not simply as part of new forms of mechanised reproduction but in relation to a broader process of *normalisation and subjection* of the observer.' (**Crary 1990 p17** *emphasis added*). The premise that there

was a standardisation of visual imagery in the nineteenth century is unproven: this was the period when the proliferation of image types which had begun in the previous century continued: cartoons, engineering drawings, architectural renderings, botanical illustrations, fine art in a wide variety of styles, photographs and pictorial advertisements. Crary's fundamental point has some value: that if one kind of image becomes regarded as 'realist' this has a polarising effect on other images, but the vocabulary of subjection seems misplaced. For Harvey 'representations of space and time arise out of the world of social practices but then become a form of *regulation of* those practices.' (**Harvey 1996 p212** *emphasis added*). Jenks (**1995 p7**) claims that 'the overwhelming appeal of such a rigid and intransigent relation between vision and visual field [that is, linear perspective] must surely derive from its strengths in protecting the variety of interests inherent in any social order of signs and images.' This invective of conspiracy seems to ignore the fact that the claim of photography (for example) to be the arbiter of realism arises from popular perception of its special relationship to the visible world, not from oppression. Who is doing the oppressing is not made clear. When Virilio (**1994** (**1988**) **p21**) claims that 'Considered irrefutable proof of the existence of an objective world, the snapshot was, in fact, the bearer of its own future ruin,' one can only wonder what form this ruin took or is taking.

Not only is Crary simply not correct in suggesting that there is a single hegemonic type of image in the nineteenth century, there is little sign of such a visual monoculture in more recent times. In this thesis it will become clear that spatialities are varied, and that diverse complex and subtle spatial 'languages' have been developed, to respond to different needs. One suspects that visual culture tends to be insensitive to the different spatial usages of visual technologies because this would militate against sweeping statements about periods and cultures. By contrast I hope to offer a view which does justice to the variety of artefacts and the subtlety of their spatial practices.

Invisibility of shared beliefs and attitudes

The 'shattering of space' which seems to occur roughly simultaneously in the sciences and the visual arts around the beginning of the twentieth century has been taken as an example of the *episteme*, a set of shared values which reflect a particular culture's biases, preferences, values and oversights (**Foucault 1970** (**1966**) *passim*), a concept on which that of visual culture is largely based. The idea that individual cultures have characteristic styles of thought is almost as old as historiography itself, but the novel idea implicit in Foucault and explicit in Kuhn's idea of the *paradigm* (**Kuhn 1970**), is the emphasis on the 'invisibility' to the participants of the assumptions which the group or culture shares. This conceptual innovation seems to have arisen in several related fields; for example in relation to cross-cultural linguistic translation Becker remarks that 'for the most part, in most cultures, knowledge of plot constraints is *unstated background knowledge*' (**Becker 1995 p30-32** *emphasis added*); Hall (**1959 p30**) suggested that 'culture hides much more than it reveals, and strangely enough what it hides, it hides most effectively from its own participants'; Barthes' concept of

mythologies (**1973** (**1957**) *passim*) reflects how what is historical, cultural and conventional comes to seem natural and inevitable; Harvey, discussing time, space and their interrelation, suggests that though clock time is a social construct, modern societies accept it as an objective fact: the sense of objectivity in a shared social construct tends to make the construct invisible (**Harvey 1996 p212**).

I make use of this concept of 'invisibility' in the present thesis, for its valuable emphasis on the difficulty of pinning down a single 'truth' about spatial perception. I shall develop the argument that some spatial languages[9] (for example those of film, comics and textual layout) can be thought of as 'mature' and that this maturity has often emerged through pragmatic experiment rather than a search for objective visual truth. I shall argue that the artifice involved in their production is concealed, appearing naturalistic, because its methods have become invisible. A culture might believe that its visual representations match a truth while nevertheless experimenting pragmatically and developing spatial representations in a partly *ad hoc* way.

The role of intention

In discussing visual artefacts of many kinds I will refer to the intentions of their makers. However, this should not be taken to mean that I impute consciously formed intentions to the individuals responsible, nor that an unlimited range of actions was available to them. It would be ridiculous for example to imagine that a thirteenth-century painter had a free choice as to whether or not to use geometric perspective as we now know it. In many cases there is no possibility of knowing what such intentions might have been, and so the intention referred to is often both unconscious and collective. In one sense this is standard art-historical practice: 'The Baroque appealed through the emotions to the widest possible audience' (**Clark 1969 p182**). It also underpins the concept of visual culture, since this assumes that societies in diverse places and times use characteristic visual artefacts which reflect their preferences, taboos, world views and so forth. Culture in part comprises views of what is possible, legitimate or correct and these views in turn have a formative effect on individual and group actions.

[9] In this thesis I occasionally use the idea of **language** (for example, the 'language of film') when describing the way in which meaning can be afforded by assembling, juxtaposing and sequencing components. I mean to imply no view as to whether this might be more than a metaphor.

3 Depth perception and depiction

1 Introduction

In this chapter and the next I deal with two problems concerning the planar representation of world scenes in still pictures.[1] These problems are (1) that what is to be represented is sometimes considered uncontroversial – this is the idea that what is 'out there' to be pictured is easily and objectively established; and (2) that there is a kind of picture which is definitively 'realistic' – the belief that particular ways of using planar media to depict the third dimension are correct in an absolute sense. Is there a pictorial representation which is fundamentally realistic in the sense that it is a correct representation of a scene? If there is, then such pictures would be *essentially* pictorial, and looking at the picture would be like looking at the scene. Any other picture would be a divergence from this archetype and correspondingly less correct. I shall argue that no such perfect representation is possible, and that the definition of realism is instead dependent on the purposes of the artefact. The unattainability of perfect realism I consider to be fundamental, and not simply a practical difficulty arising from the limitations of media.[2]

My purpose is to demonstrate that:

- there can be no still picture which fully imitates the experience of natural vision[3]
- the ways in which pictures fall short of matching vision is a vital part of their *expressivity*[4]

I therefore discuss two aspects of the relationship between natural vision and pictures: (1) issues arising from the attempt to match pictures to vision and (2) reasons why one might want *not* to match pictures to vision. For those specifying and designing pictorial information – computer graphics, virtual environments, pictorial multimedia interfaces – this last point is the most significant. Expressivity is not necessarily increased by greater realism, however defined. Nevertheless I later show that what might be called the *illusion of realism* has an important role to play in some kinds of pictures.

This chapter concentrates on the variety of means for suggesting depth in pictures, and argues that each method has its own distinctive potential. Choosing to use some depth cue rather than another leads to specific results both in terms of information conveyed and in the relationship created between the picture and the user. The work of Gibson in relation to depth *perception* is used in an analysis of depth *depiction*. This provides a taxonomy of depth cues which serves as a basis for subsequent discussion.

---

1 As indicated in the Introduction, I do not differentiate in this thesis between representations which are derived from actual scenes really observed by the picture-maker, those derived from the imagination and those which are some combination of the two.

2 Depending on the technology employed, the shortfalls of media when compared with observed scenes include the smaller tonal gamut of painting, the narrower angle of view of VR displays, the limited colour range of printing, and many others.

3 In a sense this is uncontroversial given that the pictures discussed are still, while the world and the observer are not. However the exact ways in which such pictures fail to imitate the natural vision of scenes yields important insights into the nature of picture-making.

4 This term was given an initial definition in the Introduction.

2 Definitions

In discussing the third dimension and the issues arising from its representation in planar media (treating physical surfaces and digital displays alike for reasons which will be explained) some preliminary definitions are necessary.

Depth and dimensions

Jones (**2001 p56-61**) points out that space can be defined as 'three-dimensional' in various ways.[5] When used in this thesis with reference to the real world, the term is generally used simply to denote that the world has volume. Specifically, the three dimensions are measured against Cartesian axes in relation to the observer representing distance across the field of view, distance up or down the field of view, and distance from the observer. There are cases to be made for rival systems. Lannoch and Lannoch (**1989 p41**) object that any such system little resembles our experience of the world, that 'the three-dimensional geometric model in particular fails to account for specifically human perceptions and points of view.' To reflect a view of the world as enveloping the user, it might be considered preferable to use spherical polar coordinates representing how far up or down, right or left (or a combination) the observer turns, together with the distance from the observer of the various parts of the environment, in order to see or reach some part of the scene. But here, for simplicity's sake, I use the same Cartesian terms for real space perceived and for depicted space.

In relation to depiction, I use the term three-dimensional to distinguish images appearing to have volume from those which are purely planar, and the word *depth* to denote one of those dimensions. Considering pictorial space in Cartesian terms the *plane* is taken to be that on which the representation appears, its width and height accounting for two dimensions, while *depth* refers to the remaining dimension beyond the plane[6] which (assuming that unevenness of substrate or media is ignored) only exists through some illusion or depiction.

Figure 3.01.
Schwitters, Kurt. *Circle*, 1925
For this thesis such artefacts are not considered to be pictures.
From **20th Century Art Book 1999** Phaidon, London p418

[5] For example in Cartesian coordinates, three distances; in cylindrical polar coordinates, two distances and an angle; in spherical polar coordinates, a distance and two angles. The location of a point in any of these systems may be represented in each of the others: though conceptually they suggest different models, they are equivalent.

[6] Depth can be conceived before as well as beyond the plane, and some pictures have attempted to represent it, but for this thesis depth can be considered as always beyond.

The working definition of a *picture* is that it is a representation which resembles an actual scene. The term therefore excludes strongly abstract artefacts such as the late works of Mondrian or the collages of Schwitters (Figure 3.01) which though they may exhibit apparent depth are almost certainly not to be taken as representations of scenes. Quite what the resemblance between scenes and pictures may be is a fundamental theme of the chapter. The thesis hinges on the nature of this correspondence between actual space and pictorial space, which has strong implications for how any representation is designed. In particular I demonstrate that the relationship between world space and pictorial space differs with the functions which pictorial representations are intended to serve.

Distinguishing pictures and diagrams

The argument will be made here that, though pictures represent in a different way to diagrams, the design of both is affected by pragmatic decisions about the use of space. On the face of it, this seems a strange assertion. In a schematic representation, if element P appears to the left of element Q, this need not denote that the thing signified by P is to the left of the thing signified by Q, whereas in a picture – by definition – that is just what is denoted. It would be careless therefore to suggest that the schematic and pictorial represent in just the same way: it seems that in schematic representation the significance of the spatial relation must be 'decoded', whereas in the pictorial it is 'just seen', without the intervention of any symbolic mechanism. Nevertheless, I shall develop the argument that many aspects of the space in pictures are arbitrary in the sense that they are chosen to fulfil particular purposes rather than being dictated solely by the properties of the scene and that there is a continuum which connects the schematic to the pictorial. The study of spatiality in planar media highlights the way in which all visual representations embody pragmatic decisions: the designer may benefit by recognising this.

Levels of representation

Considering a picture ***P*** of an object or scene in the world (which I will refer to as the model, ***M***) it is useful, rather than thinking only of how ***M*** is represented using picture ***P***, also to consider briefly what ***M*** itself represents. Ostensive subject matter – a landscape, a streetscene, a human subject – often has a metaphorical or symbolic as well as a literal relation to meaning so that, for example, a dove represents peace or a trash-can represents the concept of deletion. Taking the simple case of a scene which represents a single idea – such as deletion – if this is denoted as ***I***, one can begin to map the stages of representation from idea to picture in the form:

$$\boldsymbol{I \rightarrow M \rightarrow P}$$

It is important to note that the model is chosen as a vehicle for an idea: it is rare for something to be depicted for the sake of depiction.

This illuminates the problem of Mackinlay's test of expressivity (**1986 p114-6**, see Introduction p9), if the 'facts in the set' are considered to be the ideas for which the model stands, then a picture can represent *more* than the set of facts by having both a

simple pictorial and a metaphorical relation to its subject. In addition, any picture presents *less* than is in the 'model' since few representations can show all aspects of the model, particularly when a three-dimensional world must be mapped to the plane.

In most cases the idea ***I*** is selected to serve some objective ***O***. This could be diagrammed as:

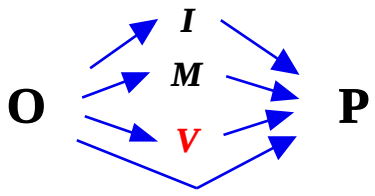
$$O \rightarrow I \rightarrow M \rightarrow P$$

This needs to be further developed, to represent the fact that another representational process takes place between the model ***M*** and the picture ***P***. The model ***M*** is pre-pictorial and probably three-dimensional. The picture ***P*** is flat. The stage in between is the one with which this chapter is primarily concerned. It is useful to borrow here from the pipeline approach of synthetic computer graphics (**Lansdown 1987 p54-55**, **Foley *et al* 1995 p334-5, 806-9**), in which it is normal to conceptualise the difference between two sets of decisions in the design of a picture: one involves the geometry and other attributes of the model ***M*** – that which is to be depicted – while the other relates to the visualisation of the model necessary to its display, including such factors as projection system, mode of rendering, point of view, framing and so forth. Both are processes of representation. Denoting the viewing parameters by ***V***, the stages can be redrawn as:

$$O \rightarrow I \rightarrow M \rightarrow V \rightarrow P$$

In sum, this crudely represents how the purpose or objective ***O*** informs the selection of an idea ***I*** which is instantiated in some particular model ***M***. This in turn is viewed ***V*** in a particular way, and also pictured ***P***. It should be noted that this staged diagram is equally applicable whether the picture are traditional or digital, static or moving, highly realistic (eg. photographs) or more abstract (eg. 'icons' in computer interfaces). For simplicity the diagram shows each of the stages as singular. In fact, objectives are likely to be many for a single picture, while a single model may yield many views and many pictures, and so forth. It is also shown as linear, when in fact, as discussed later in the thesis, it might also show various forms of 'feedback'.

While adequate as a pipeline view of the process, the diagram could be read as implying that the effects of the objective ***O*** are confined to the next stage (the selection of the idea ***I***) and so on**.** It is part of the purpose of this thesis to show that the objectives of a representation are best considered as strongly influencing *all* these processes, including the viewing stage ***V*** and even the depiction itself ***P***. As a map of the influences at work therefore, the following is perhaps a better representation.



The inevitability of depiction

As already indicated, models do not themselves exist visually: it is only rendering which makes them visible, and, except where actual physical three-dimensional artefacts are made, all representations to be received via the eye must be made pictorial. The technology used to achieve this is irrelevant to the general principle. It makes no difference whether the user looks at a single planar image on a monitor, two such images in a headmounted display, or even has the image beamed directly onto the retinae (in which case the planar images are virtual but are still identifiable in the visualisation process). Because each retina is responsive only to a two-dimensional array of light, ambiguities are acknowledged to permeate depth perception. Gregory (**1970 p25**) refers to the 'infinity of possible three-dimensional shapes' and Hoffman (**1998 p13**) describes as the fundamental problem of vision the fact that 'the image at the eye has countless possible interpretations.' In *some* ways, therefore, how observers interpret a picture is analogous to how the retinal image itself is interpreted.[7] While vision may be supplemented by background knowledge and by the evidence of other senses, it is essentially a pictorial mechanism in the sense that the eye is a visual device only and one which itself deals with two-dimensional projections, not three-dimensional information. To stimulate this device using planar media in order to suggest the appearance of some scene, a process of picture-making is required: there is no escaping the need to map the three-dimensional world to a surface.

3 Depth perception and depiction

Some aspects of depth perception are as fundamental and irresistible as gestalt principles, and indeed one gestalt principle is itself concerned with depth perception: relative size is shown to promote the differentiation of figure and ground.[8] Colour and tone, focus and other factors, can all provoke a sense of depth in the absence of ostensive depiction. Mondrian, having eliminated subject matter from his paintings but having elected to use the psychological primaries red, yellow and blue, found that he could not get his paintings to look as flat as he wished,[9] and began to enclose the panes of colour in a black grid in order to reassert the sense of flatness (**Oxenaar 1982 p75-6**). Depth insisted on appearing even in such highly abstract subject matter.

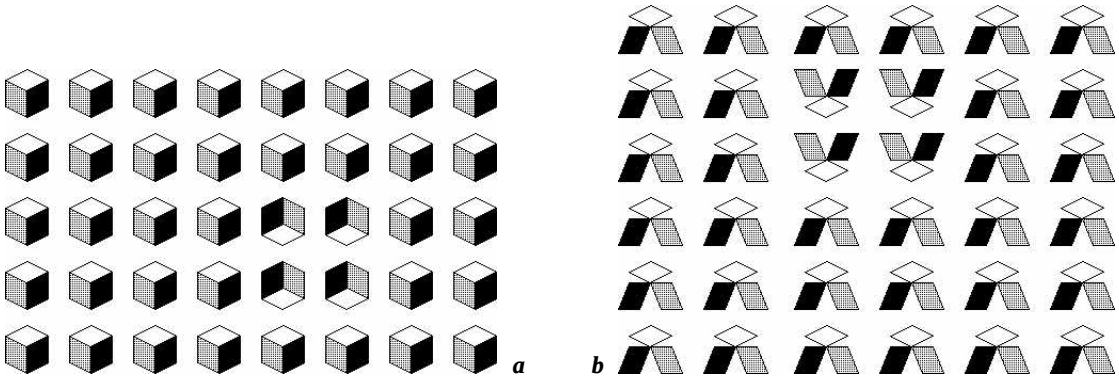


Figure 3.02. Taken from Healey 1999

In **a**, shapes are combined to represent three-dimensional solids; in **b** similar shapes are combined to confound any depth impression. Within a pre-attentive timescale (less than 200 to 250msec) the zone of difference can be detected in **a**, while in **b** it cannot.

[7] I discuss later whether the curvature of the retinae is significant.
[8] As always, conflicting factors can undermine this principle.
[9] The inherently spatial properties of colour are discussed later in the chapter.

Enns and Rensink (1990 cited in **Healey 1999**) seem to have shown that three-dimensionality can enable even pre-attentive perception of differences in a scene (Figure 3.02a-b). This also suggests that depth-awareness operates at a very low level, before mental processing has had time to recognise the shapes as depictions of cubes.

It is hardly possible to make pictures without suggesting depth, at least of the figure-ground type. The difficulty is aggravated by prior knowledge in that, as soon as something in a picture is recognised, the viewer tends to impose on it a depth awareness which has been learned from the experience of similar objects in the real world and of other pictures. Perception of depth is thus in part informed by world knowledge. For example the perception that a line of telegraph poles recedes may be assisted by the assumption that such poles are normally of uniform height.[10] However, many pictures have been made which suppress or minimise particular depth effects, and such images form part of the discussion of the relation between depth cues and picturing below. I will argue that selectivity over the depth cues employed is a vital aspect of expressivity.

Depth perception is often attributed primarily to binocular vision, and it is certainly worth considering that the overlap of the visual fields of the two eyes necessary to stereopsis has been achieved at the price of losing the panoramic vision possessed by animals with side-facing eyes. The work of JJ Gibson (**1950**, **1979**) is important in emphasising the range of stimuli evoking spatial perception which are *not* dependent on binocular vision – and which are therefore available for use in single planar images. In fact, of Gibson's 13 stimuli (or quasi-stimuli: he calls five of them probable signs) only 2 require binocular vision. This is not to deny that for a complete spatial experience binocular stimuli are necessary, but does help to indicate the range of other stimuli which can be used in monocular artefacts. It also may help to explain why we are generally satisfied with the monocular spatiality of images such as photographs and films. Similarly, though movement of objects or scenes relative to the observer is of fundamental importance to depth perception in natural vision, only two of Gibson's cues are directly dependent on it.

There is an obvious irony in applying Gibson's pictureless taxonomy to pictorial images. However, Gibson's work serves three purposes here: (1) it provides a vocabulary of terms useful in subsequent discussion; (2) it has the standard taxonomic advantages of identifying individual issues and disentangling them from generalised approaches and (3) it highlights the similarities and differences between looking at pictures and the perception of the natural world. This last is a particular reason for choosing Gibson's approach rather than any other, since he explicitly sets out to systematise our understanding of the perception of the world without resorting to pictorial concepts.

[10] Another possible determinant is the principle of minimal assumption, in that a set of similar objects may be assumed to be the same height unless there is contrary evidence. Hoffman proposes a set of rules to which he claims vision conforms which are mostly variants on this idea of minimal assumption (**Hoffman 1998**).

Gibsonian depth cues

Gibson's 'varieties of perspective' classifies postulated cues into groups (**Gibson 1950 p137-144**).

Group A – perspectives of position
1 Texture perspective
2 Size perspective
3 Linear perspective

Group B – perspectives of parallax
4 Binocular perspective (requires binocular vision)
5 Motion perspective (requires scene / observer motion)

Group C – perspectives independent of observer's motion or position
6 Aerial perspective
7 Perspective of blur
8 Relative upward location in the visual field

Group D – perspectives perceived at contours (edges)
9 Sudden shift of texture density or spacing
10 Shift in amount of double imagery (requires binocular vision)
11 Shift in the rate of motion (requires scene / observer motion)
12 Completeness or continuity of outline
13 Transitions between light and shade

It is immediately apparent that a traditionally important set of cues is missing – shading and cast shadows. In fact, Gibson breaks his own classification scheme by including graduated shade variation in cue 13, within the group which is supposed to be confined to edge phenomena. Shadow perception may belong partly to very early stages or very low levels of perception, but Baxandall (**1995 p40-1**) remarks that shadow perception 'might well be considered as a systematic activity [ie. not at stimulus level] in the sense that any one shadow needs to be established within a larger pattern to signify with any force: a solitary, uncoordinated, uncaused shadow may be just a dark patch.' Shadow is an aspect of depth perception which has historically been extensively studied by artists with the specific intention of constructing spatial illusions. Because of this, I have introduced a fuller discussion than Gibson would have approved under his final category.

Nearly all of the examples described here are paintings, for the simple reason that taken together they represent a far wider range of approaches to depth depiction than any other medium. Media with a more limited gamut of effects such as ink drawings are unable to exploit depth cues such as colour. Photographs on the other hand tend to be unable to omit or suppress size and linear perspectives. A 'hand-made' medium such as painting offers more examples of selectivity and pragmatic decision-making than the general run of algorithmic images. Nevertheless it is impossible in most cases to offer a pictorial image which corresponds to one Gibsonian cue and to that alone; for this discussion images have been chosen which use a subset of the full range of cues, perhaps prioritising one.

Texture perspective

There is a gradual increase in the density of texture of a surface as it recedes into the distance.

Figure 3.03.
Van Gogh: The Café Terrace on the Place du Forum at Night, September 1888
From Uhde, Wilhelm 1972 *Van Gogh*, Encyclopaedia Britannica, London, Plate 23.

Van Gogh regularly used texture perspective. In some of his drawings it is an indispensable cue (Figure 3.04.). In Figure 3.03, geometric perspectival construction (Gibson No.3) is also present, but its effects are reduced for example by the way in which the strong slope of the orthogonal of the blue shop front is similar to that of the transverse canopy. The texture of the cobbles therefore plays a significant part in constructing the space. Similarly in the drawing the convergent lines of field edges are made less decisive by the odd angles of other field boundaries so that more of the work has to be done by texture.

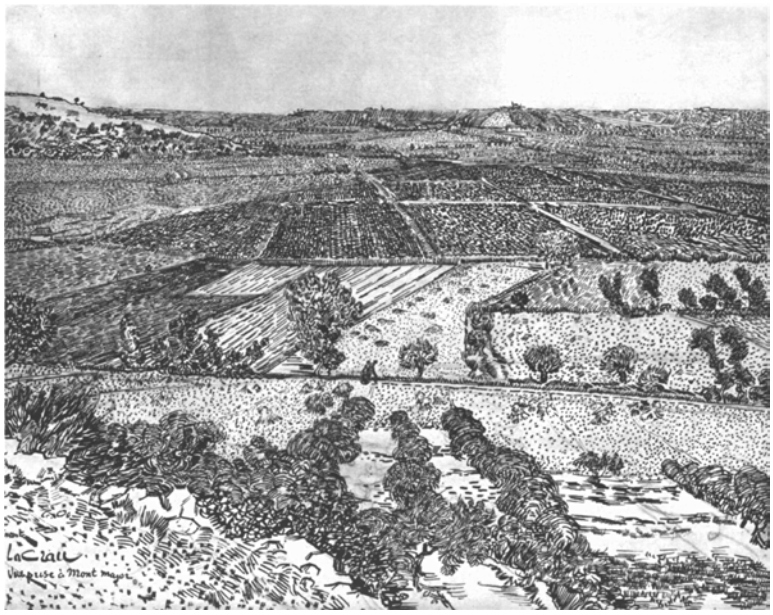


Figure 3.04.
Van Gogh: The Plain of La Crau 1888.
From Uhde, Wilhelm 1936 *The Life and Work of Van Gogh*, Phaidon, Vienna, Plate 29.

As with many of the Gibsonian cues, the presence of differential textures is not in itself diagnostic: an alternative explanation for an apparent change in texture may be that the texture-giving components are themselves variable. The default assumption is likely to be of uniform texture, however.

Texture has been one of the more recent preoccupations of the computer graphics industries. Generally it seems that the recent growth in the use of texturing has been

particularly with mimesis in mind: little work has been done on the use of textures in more abstract environments. Popular mainstream artefacts include entirely synthetic films such as Toy Story (**1995**) and Toy Story 2 (**1999**), where considerable time and expertise is invested in making detail 'bumpier and dirtier' with the specific aim that it should not 'look like it was made on a computer' (**Toy Story 2000 p29**). But, following Gibson, an ancillary advantage of texture would be its enhancement of the sense of depth.

Size perspective

Figure 3.05.
Paul Nash: We are Making a New World, 1918.
From Leeds City Art Galleries, 1988 *100 Years of Art in Britain – an exhibition to celebrate the centenary of Leeds City Art Gallery*, p34

For a set of similar objects, decrease in size will be interpreted as greater distance.[11] Factors at work in Figure 3.05. include discontinuities of contour in the occlusion of the profiles of the earth ridges by trees (Gibson No.12) and, to a small extent, changes in tonal contrast (No. 6 aerial perspective), but a key mode of depth articulation is relative size. In the central band of the picture, unusually little contribution is made by relative upward location in the visual field, since the full height of the distant tree stumps is seen between the nearer ones: scale dominates here.

There seems to be some evidence from vision studies of subjective perception of distant objects such that they appear larger to the observer than they should if considered purely optically. This may begin to suggest that there is some divergence between truth to the optical fact (as defined for example by measuring the size of objects as cast on the retina) and truth to perception more subjectively conceived. Such issues arise often in this chapter and inform much of the ensuing discussion of realism.

In a counter-example (Figure 3.06), the principal determinant of size is not distance but social position and dramatic importance. King Herod is apparently at the back of the scene, according to the interruption of contours and height in the image field, but is larger than the figures who are apparently nearer. Such anti-perspectival scaling is of course common today in *diagrammatic* graphics where visual mimesis is not an overriding concern. This image is certainly not devoid of spatial depth, which has

[11] Texture and size perspectives might be considered as the same thing. Gibson (**1950**) distinguishes them on the grounds of how they are captured in perception, while Marr (**1982 p233-9**) and Gregory (**1998 p190-1**) are more wary of this distinction.

Figure 3.06.
Stained glass window, The Massacre of the Innocents, St Peter Mancroft Church, Norwich, UK , 15th Century.
From Lee, Lawrence; Seddon, George and Stephens, Francis 1982 *Stained Glass* Mitchell Beazley, UK, p26.

been used effectively to condense a large amount of incident into a small space. This condensing of the image also facilitates rhetorical juxtaposition (such as the sword-pierced baby between the heads of Herod and a horrified parent).

Figure 3.07.
Crivelli: *The Annunciation with Saint Emidius*, 1486.
From **Cole 1992** p23

A number of other depth cues are notable by their absence, including atmospheric and textural transitions (Gibson Nos. 6, 1, 9). What linear perspective (No.3) there is undermines the depth-wise spatial coherence of the whole, since the orthogonal of the bed points away out of the picture rather than into it. However, in two-dimensional terms it helps to frame the scene.

Even when there is an apparently rigorous application of the rules of geometric perspective, all is not necessarily what it seems. Hart and Robson (**1999**) have 'reverse-engineered' a computer model of the buildings and figures depicted in *The Annunciation with Saint Emidius* (Figure 3.07) and shown many deliberate deviations from size perspective, in the interests, in their view, of reconciling depth with intimacy. The range of reasons for such 'subversion' of 'correct' perspective are discussed later.

Linear perspective

In picture-making there are many varieties of linear perspective or projection systems, whereas in the study of vision, including in Gibson, there is usually only one, that of convergent linear perspective. This suggests that the convergent representation is correct, since it is based on vision, and that others are wrong or imperfect. The major issues raised by such a supposition are discussed in the next chapter; here I offer examples of various kinds of apparent rule-breaking in the making of pictures, and show the range of motives which these practices might serve.

Non-convergent systems include the orthographic, axonometric, and isometric (all varieties of parallel projection). In the isometric all lines are drawn to scale, rather than decreasing in size with distance, the cardinal horizontals are inclined at 30°, and all planes are equally distorted. Approximations to such views have been used in Roman, Byzantine, Persian and Chinese paintings and extensively in Japanese woodcuts (**Dubery and Willats 1983 p38-9**). In axonometric projection horizontal surfaces are drawn in rotated plan view, and the necessary verticals and horizontals are then appended to them (***op cit* p28-9**). This projection has attracted architects throughout the twentieth century, because of its ability to combine an undistorted plan with an evocation of the character of the interrelated spaces.

Figure 3.08.
Uccello: The Battle of San Romano (c1450s).
From **Cole 1992** p16-17

Since the Renaissance, when the rules of perspective construction in painting were developed and codified by Alberti, Viator and Dürer and others (**Ivins (1975) 1938**), linear perspective has become so identified with the depiction of depth that it is often used as a synonym for it, though Gibson draws attention to the fact that it is only one cue among several. He also points out that its full effectiveness is contingent on the choice of subject matter. In the example illustrated (Figure 3.08.) Uccello has engineered the positioning of the armaments on the ground to create an improbable number of orthogonals. This is a particularly obvious example of a picture-maker contriving a relationship between subject matter, viewpoint and perspective cues, but this threefold synthesis later becomes more subtle and effective as the use of convergent linear perspective develops into a mature spatial practice. In Uccello's time perspective was still an 'added feature' rather as it might be in a computer game today. Considerable aura was attached to Uccello's technological prowess in its own right. Elkins suggests that, for most Renaissance artists, perspective was *in paintings* – rather than paintings being in perspective. Up to and beyond the time of Vasari, perspective treatises enumerated the perspectival parts of pictures such that a good painting could be 'full of perspectives' (Vasari quoted by **Elkins 1994 p55**).

The Uccello painting shows the use of two combined projections: the landscape background does not conform to the same perspective as the foreground. Solso (**1994 p160**) suggests that it was a lack of competence that led Uccello to use two different perspectives, but given what is known of Uccello's general mastery of perspectival construction this is hardly likely. The more probable explanation is that he wanted to construct a relatively flat background in order to afford two 'incompatible' views within one picture, so that the distant scene is viewed optimally and so is the near one. I will show in the chapter on Film that this contriving of *optimal viewpoints* becomes a decisive influence on the construction and use of space in film. The flatness of the backdrop also serves to bring the viewer's attention back to the main subject of the painting, avoiding the 'hole in the wall' effect which a powerful perspective recession would have produced. Such combined motives are probably more common than the pure case of a particular spatial device being selected for a single reason.

Figure 3.09.
Active Worlds: Scene in Active Worlds 2.2.

In case such perspectival hybrids seem to be a thing of the past, it is worth noting how a simple VR browser like Active Worlds uses two projection systems in a rather

similar way (Figure 3.09). The distant landscape scene is a 2-D picture mapped to a vertical cylindrical surface. It is always at the same distance however closely the user tries to approach it. The near environment in which the user moves employs a conventional 3-point perspective projection of real 3-D data.

A pictorial map of the Piccadilly Line of the London Underground (not illustrated) makes an interesting contrast with the Uccello painting. Though a fanciful rendition, it conforms to a single unified perspective. The railway is viewed from the western central area of London looking east: under these circumstances the rules of convergent perspective have a somewhat similar effect to the non-linear scaling of the Beck map (**Garland 1994**), enlarging Central London distances and diminishing those in the distant outer region. However, the continuous nature of such a unified perspective means that the most distant stations are entirely lost in the horizon. Only by some degree of 'flattening' of the more distant terrain, as in the Uccello, could these distant regions have been adequately depicted.[12]

Figure 3.10.
De Chirico: The Delights of the Poet, 1913.
From Rubin, William 1982 *De Chirico* Museum of Modern Art, New York, p144

Once the 'rules' of linear perspective have been widely acknowledged, they can be subverted. The De Chirico painting (Figure 3.10.) relies on the viewer's prior knowledge, not only of the world (for example in assuming that the basin is rectangular and that it is parallel to the buildings of the square) but also of picture-making, in order to disconcert. A recent virtual environment project has attempted to build a virtual city 'based on locations seen in de Chirico's works, in which a user can freely navigate, simulating the experience of actually walking round the places imagined by the artist' (**Druks 1995 p113**). Druks hopes that 'a faithful rendition could allow the user full access to the world the artist could only metaphorically imply access to.' He seems oblivious to the parodic nature of de Chirico's imagery, which relies on the abuse of standard perspectival conventions to offer the user a difficult choice: either this is a straightforward depiction of an unconventional world, or it is a distorted vision of conventional reality (or it is a mixture of both). Either way, it is self-consciously *about* painting, amongst other things: the digital project shows a fundamental misunderstanding of the nature of picture-making.

[12] Tufte (**1990 p12-3**) shows an unusual Japanese solution to this problem in which a pictorial map gives way at one edge to a highly compressed schematic one, so that not only the 'perspective' but the mode of representation changes abruptly.

This painting makes an interesting contrast with the Renaissance images which it parodies, particularly in the way that the viewer is invited to consider what is to left and right of the scene shown. The prototypes of this image were built on centrality, even to the extent – in early examples – of using only a central vanishing point, but here the odd angle at the foot of the arcade to the left invites the viewer to imagine looking to that side, while the enormous shadow cast by something which is invisible to the right acts as a similar unresolvable invitation. The invitation and the impossibility of its resolution illustrate the way in which authorial prohibition may be a vital element of pictorial expressivity, prefiguring the discussion of such techniques in film.

Figure 3.11
Suzuki Harunobu: Woman with a fan at the garden fence, 1766-70

An example of the non-application of linear perspective is this Japanese print (Figure 3.11.). Not only do parallel lines in the scene not converge in the image, but the whole articulation of the space is at odds with the western tradition. This is another example of the synthesis of a linear projection system with a particular viewpoint. The orientation of the subject matter together with the parallel projection scheme constitute what in Western draughtsmanship would be called an isometric view. Far from receding into a vortex-like central vanishing point, the scenery seems to rush out of the sides of the picture. The surface is simultaneously flat and deep in interesting ways. One effect of the meeting of the two oblique planes behind the figure is that the woman seems to be pushed forward, nearer to the picture plane than the position of her feet would suggest. The depths in the space are not so much an imitation of depths in the world, but a making of depth in the picture's own terms. This is, in Gombrich's phrase, 'the dominance of making over matching' (**Gombrich 1977 p248**) which recurs throughout this thesis.

Again considering the Gibsonian cues as a palette of possible effects, it is clear that some are used very strongly in this image, while others are omitted. One can see where the myth of the flatness of Japanese art comes from: there is no modelling with light and shade within surfaces. In addition, texture perspective and size perspective are absent. There is only a minimal suggestion of aerial perspective since the local colour of objects is used. However this local colour is assigned to parts of the scene in such a way that the nearer parts are warmer in colour than the further: again, this is as much a tactic to *make* depth as to depict it.

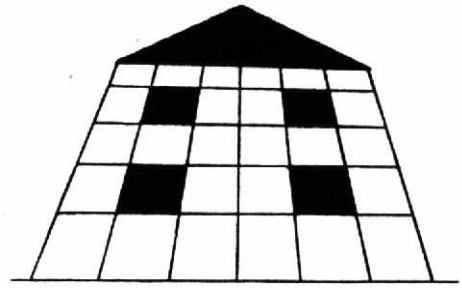


Figure 3.12.
Klee: *Figure 44: Once more the Vertical.*
From **Klee 1925/1968** p41

In relation to the house-front depicted in his Pedagogical Sketchbook (Figure 3.12) Klee asks:

> Why is Fig 44 as representation of a house wall incorrect? It isn't wrong logically. The lower window openings are closer to the eye than the upper ones, which means they are larger perspectively. As representation of a floor pattern, this perspective rendering could be easily accepted. This picture therefore is not incorrect **logically**, but **psychologically**. Because every creature, in order to preserve his balance, insists on seeing actual verticals projected as such.[13]
>
> **Klee 1968** (**1925**) p41

Similarly Gregory (**1977 p174**) suggests that 'aiming a camera upwards, to take in a tall building, gives the impression of the building falling backwards. And yet this is true perspective.' For Gombrich, writing in 1980, such upward views onto buildings are accepted 'without demur' whereas he admits to finding the downward equivalent (Figure 3.13) 'a trifle less easy to accept' (**Gombrich 1980 p187**).

Figure 3.13.
Rockefeller Centre from top of Time-Life Building, from Andreas Feininger's The Face of New York (New York 1954). Photograph.
From **Gombrich 1980** p189.

[13] The difficulty of this particular image is probably aggravated by the need for the viewing station to be below ground level to yield such a view. However, this is not the whole explanation, as Gregory's and Gombrich's examples make clear.

The three comments provide an insight into the continuing development of pictorial conventions. What was unacceptable in 1925 or even in 1977 is probably widely regarded as natural now. Unlike Klee and Gregory, Gombrich is aware that this is symptomatic of changing attitudes to the 'correct' brought about by exposure to new kinds of images. Snyder (**1980 p232-4**) argues that the design of photographic apparatus was guided in the early decades by the example of painting, specifically with a view to eliminating 'distortion,' but it seems that now the photograph is the arbiter of the real. We have become so accustomed to the convergent perspective of photographs taken looking up or down rectilinear structures that our perception of what is normal has been modified.

The lack of logical consistency in the former convention is certainly remarkable: verticals in a scene are to be made verticals in the corresponding picture on the grounds that that are they are *known* to be vertical in real life, but the same does not apply to horizontals, where the whole point of linear perspective is to ensure that horizontals are aligned in accordance with the optical image, *not* in accordance with what is known. This conditional visual 'realism' is the one to which many kinds of pictures conformed from the Renaissance until recently. Though subsequent developments have made Klee's view seem rather archaic, his distinction between the *logically* and *psychologically* correct is still a vital one to which I return later.

Binocular perspective

The separation of the two eyes means that each receives a different image. Binocular disparity is much more apparent at close range (5m or less) than for distance viewing, since it decreases in proportion to the square of the distance (**Bruce, Green and Georgeson 1996 p140**).

Mainstream perspectival painting is notoriously monocular. Brunelleschi's demonstration[14] is based on viewing with a single eye, and the history of painting and photography has continued in this vein with rare exceptions. Only by using two images and making each available to only one eye, can the effect of binocular perspective be simulated. The technique has been applied to photography, film, drawings and schematics such as stereograms, and is the basis of most VR systems.

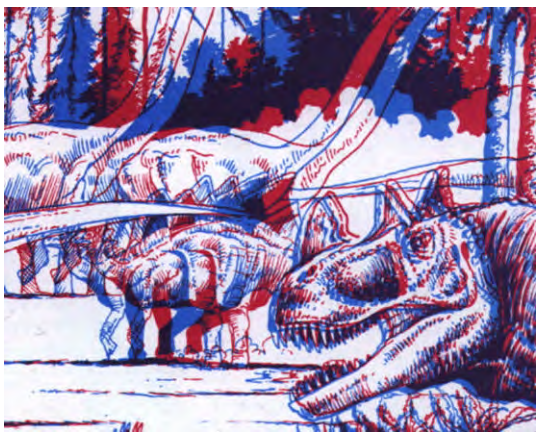


Figure 3.14.
Stereoscopic drawing for viewing through a red and green pair of gels (detail).
From *The 3-D Book of Dinosaurs* Orbis Publishing 1998 p241

[14] In 1425 Brunelleschi made a picture (now lost) of the Florence Baptistery which was allegedly remarkable for its match to the actual scene. Issues raised by this are discussed in the next chapter.

The heyday of stereoscopy in still pictures belongs to the 1850s. In 1833 Wheatstone proposed that the mental fusion of the images from the two eyes was the cause of spatial perception (**Gautrand 1998b**) and created the stereoscope to demonstrate that this was so. Wheatstone's device used drawn graphics, but Brewster combined Wheatstone's invention with photography in 1844 (***ibid***). A pair of photographs could be made with two adjacent camera lenses to record a stereoscopic view. By the mid 1850s, one version of the stereoscope had sold more than a million in England alone, and in 1851 one was exhibited in the Great Exhibition, where it was patronised by Albert and Victoria. By the end of the decade the ambition of the London Stereoscope Company – 'no home without a stereoscope' – was almost fulfilled (**Macdonald 1979 p50**). Various exotica were available, ghost pictures, moral tableaux, freaks and oddities, and pornography (Figure 3.15).

Figure 3.15.
Colombier: Pornographic stereograph c. 1862.
From McCauley, Elizabeth Anne 1994 *Industrial Madness: commercial photography in Paris 1848-1871*, Yale University Press, New Haven p179

In general the motivation for using stereographic images seems to have been affective rather than informational. In both Figures 3.14 and 3.15 it is the 'sense of the real,' or the suppression of the unreal, which is the principal objective. In terms of the informational expressiveness of static pictures, there is remarkably little more value in a binocular view than can be got from the corresponding monocular picture, but what is gained is a sense of co-presence. In the case of both these figures it is easy to see why this is an important objective.

Only relatively recently have other applications emerged in which binocularity is combined with motion, such as the use of virtual environments for aeroplane or oil-rig simulators, distance surgery and architectural visualisation. Here the information function is greater, though it is probably still the case that most of the informational benefits come from motion rather than stereopsis.

Thwaites (**Thwaites 1999 p222**) traces 'our 1990s fascination with creating the third dimension in two-dimensional space' from the prehistory of painting, via early

stereoscopes, to virtual reality, but this seems to imply a continuum from the Victorian stereoscope to stereoscopic virtual environments which is not borne out by the facts. After the third quarter of the nineteenth century the stereograph was relegated to the status of a child's toy. Its ability to present the third dimension seems somehow to have become superfluous. What might have led to this demise, while straightforward photography continued to increase? The great attraction of photography lies in its verisimilitude, yet stereopsis becomes neglected. Stereoscopes were available quite cheaply, and many were originally bought. Perhaps a means of communication which required a mechanical device for viewing could not be assimilated into normal life; but the gramophone and the television, both dependent on physical machinery, were adopted on a continuing basis. A deduction from this decay of the third dimension might be that there is not one realism but several, that each has its own concomitant spatialities, and that the different realisms serve different objectives. Certainly it seems that the realism generally expected of photography is selective and is not expected to amount to full mimicry of the spatial perception of the natural world.[15] This suggests that there are *kinds* of realism quite as much as *degrees* of realism and that for different objectives, technologies and contexts varying kinds of realism are taken to be 'realistic' in a rather general, undefined but clearly contingent way. These suggestions are pursued in the next chapter.

Motion perspective

There is apparently little to be said about motion perspective when discussing static planar images, except perhaps to clarify the nature of the movement to which Gibson was referring. However, the aspect of *time* in relation to static images turns out to be of fundamental importance and is discussed later. Aspects of motion perspective are also pursued in the chapter on Film.

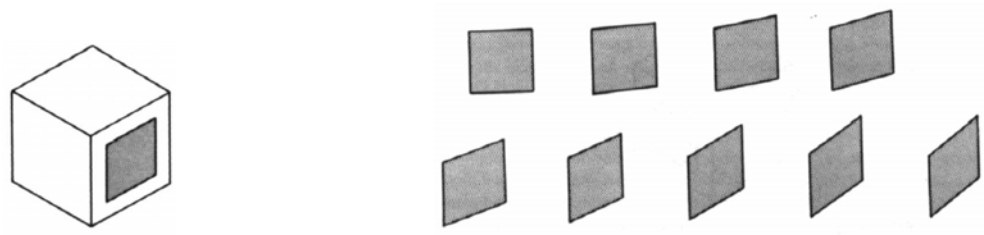
The two kinds of movement to be distinguished are:

- saccadic movement of the eye which makes it able to gather optical information: as the visual system apprehends a scene, the eye is only stationary for brief moments, so that a completely stationary eye is effectively blind. Perception is *not* well conceptualised as the aggregation of a series of static moments. 'Space cannot be apprehended except in time [...] The process of perception cannot rest on the image of a single fixation such as yields a momentary visual field. To see more than this takes time, and requires a succession of visual images. The product of these successive impressions, however, is such that, paradoxically, all awareness of the succession has been lost' (**Gibson 1950 p157**). This is one of the key ways in which vision and pictures are now acknowledged to be dissimilar. This kind of motion is the only one necessary for the depth cues enumerated so far, and for most of those still to come.
- relative movement between the observer and the scene. This is the basis of *motion perspective*. One of Gibson's innovations was the concept of the 'optic flow' which yields substantial information about movement or stasis within a spatial environment, direction of movement, orientation of movement, and changes to any of these (**Gibson 1979 p227-229**). Such motion perception has ecological advantages.

[15] '3-D' moving images have also been notable for their failure to achieve more than short term interest among audiences, despite continuous technical development since around 1900 (**Katz 1994**)

Lansdown points out that 'it is often more important to detect immediately that something has moved rather than to know straight away what that something is (or even in which precise direction it has moved)' (**Boyd Davis, Lansdown, and Huxor 1997 p21**). Similarly Gombrich notes that to know in crossing a road where every car is, but not its direction or speed, would mean the observer did not survive long (**Gombrich 1973 p228**).

While motion is an important part of the impression of depth in the real world, it seems that it should be the enemy of linear perspective in pictures, since for such perspective to be optically correct the viewer must look with a single eye and from a single optimal fixed point.[16] If the viewer moves, the illusion breaks. However, movement in relation to a picture, and the adoption of non-optimal viewing stations, is surprisingly unperturbing. Even at the height of the Renaissance it seem that there was a rather surprising tolerance of such dangers. Work using computer models by Hart and Day (**1995**) has demonstrated that there was no optimal viewing seat in the famous perspectival theatre of Sebastiano Serlio – this position was occupied by a staircase. Viewing positions which are sub-optimal and subject to change through motion are familiar in everyday experience from looking at pictures in books, on advertising hoardings or when watching television – indeed in most picture-viewing.



a **b**

Figure 3.16a-b.
Subject are asked to choose from **b** the shape they had been shown in drawing **a**. They generally pick a trapezoid more square than the correct one.
From Deregowski 1984 p77-8.

Deregowski (**1984 p76-8**) reports that when given a drawing such as Figure 3.16a and asked to pick the corresponding shape from Figure 3.16b, subjects generally pick a trapezoid more square than the correct one, as though they were mentally normalising the distortion towards a more full-face view under the influence of the fact that the box in **a** is meant to be seen as viewed obliquely. Gombrich (**1973 p230-1**) offers a related example, emphasising that it is the internal relationships within a picture as much as their relation to the viewer's position which construct a space: the road receding into a landscape still recedes, even when the picture is viewed from the 'wrong' position.

Together these observations suggest that it is not objective optical correctness which is significant for the viewer, since any image whose geometry diverged from the optically correct would presumably fail to make coherent spatial sense, but this is clearly not the case. It seems that viewers can effectively compensate for non-optimal

[16] The questions about vision and pictures which this concept of 'optical correctness' raises are dealt with in full in the next chapter.

viewing positions and see the depicted scene as though it were viewed under better conditions.

Aerial perspective

Gibson asserts that aerial perspective belongs to a family of cues which is independent of both the position and motion of the observer. In fact, as I note below, this is not strictly true, but for the purposes of picture-making Gibson's assumption has generally been followed.

According to Gibson (**1950 p141**) with increasing distance 'there is an increase in haziness, blueness and desaturation of colours.' Dunning (**1991 p43-54**), approaching the issue from the point of view of painting, prefers to split this category into two: atmospheric perspective (dependent on tonal value) and colour perspective (dependent on hue). The distinction goes back at least as far as Leonardo. Dunning considers atmospheric perspective to have been 'the most effective method of creating an illusion of depth for at least the last five hundred years,' appearing as it does in Ottoman, Renaissance, Impressionist and modern works. He remarks that it may have had a wider appeal as a technique because, in his opinion, it does not violate the flatness of the picture plane to the extent that linear methods do; this is a reminder that the 'destruction' of the picture plane is not universally considered an objective of picture-making, that a picture is often meant to be seen as a substantive object not simply as a means of viewing a depicted world.

Dunning also offers the observation that, unlike linear perspective, atmospheric perspective can be subjected to great exaggeration – often without the viewer noticing. The viewer will perceive the depth, but not the method. This suggests a use of spatial cues in order to *construct* a space, rather than to depict it. A cue which 'works' is used in a way unwarranted by perception of actual scenes and, provided the impression of the natural is maintained, the illicit encoding remains unnoticed.[17] This is a typical spatial usage which reveals a lot about the artifice of picture-making and how – for some kinds of image at least – that artifice must be unobserved by the viewer.

Cutting across Gibsonian categories 6 and 7 (see next), Dunning conflates with the tonal effect of atmosphere its effect on sharpness and on apparent volume. He proposes that recessive features are: areas with weak contrast (especially contrast with the background); loss of detail and definition of edges; and flatness (lack of modelling). Salience on the other hand is promoted by objects having strong contrasts, sharp edges and detail, and a sense of volume.

Heidrich et al. (**1999 p130**) point out that in a geometric perspective representation, even aerial perspective is, strictly speaking, subject to geometry. While simple 'fog' systems in computer graphics use the z-coordinate (the world-distance of objects

[17] It should be noted that photography cannot be a yardstick of realism in aerial effects since the lenses, filters, film stock and printing paper all significantly affect the degree of atmospheric perspective in the finished photograph. Such variables are discussed in the chapters on Screen space.

from the picture plane, not their distance from the eye) as the input to the fog-density function, these underestimate the distance and therefore the amount of fog for points on the periphery of the image. Ironically, such sophisticated renderings of atmosphere bring us back to the compulsory central viewing station of geometric perspective which the simple z-depth atmosphere avoids (Figure 3.17a-b).

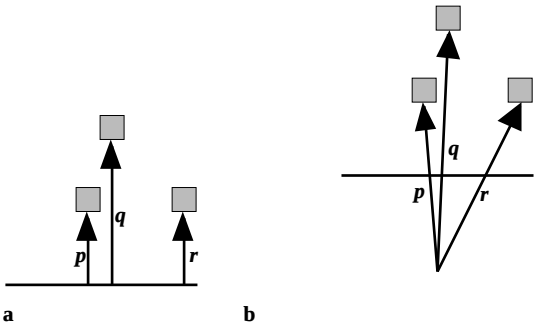


Figure 3.17a-b
Calculation of atmospheric attenuation as a function of distance. In **a,** using z-depths, the visibility of the objects at distances ***p*** and ***r*** will be equal. In **b**, using eye-to-object distances, the object at distance ***p*** will be less obscured than that at distance ***r***.
Based on **Heidrich *et al* 1999**.

Figure 3.18
Watteau: The Island of Cythera, 1709 (detail).
From Posner, Donald 1984 *Antoine Watteau* Weidenfeld and Nicholson, London, p86

Like atmospheric perspective, colour perspective can be usefully exaggerated. Dunning suggests (**1991 p47**) that this helps compensate for the lack of binocular vision and movement parallax. This depth phenomenon has a number of causes: increasing blueness with distance caused by greater scattering of blue light than red; decreasing saturation of colour caused by the intervention of atmospheric particles and moisture; and chromatic aberration, the physiological tendency for the eye to focus on red as though it were near and on blue as though it were farther away causing the visual system to misinterpret difference of hue as that of distance (***op cit* p52**).

In the Hiroshige print (Figure 3.19) both atmospheric and colour perspective are used to the full. Strong recession of the distant riverbank is provided by weakening of contrast with the background, loss of detail and lack of modelling. In addition to the spatial dynamics within the scene, a spatial effect is also provided by the strong red of the overlaid text panels. Atmospheric and colour perspective do the work which is not done by linear perspective.

Figure 3.19
Hiroshige: Shower at O-Hashi Bridge, 1857.
From a book on Japanese prints (in Japanese, title unknown), Kodansha 1970, plate 134

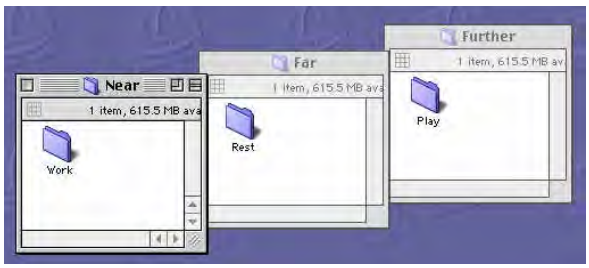


Figure 3.20
Apple Computer: Windows of the Macintosh 'desktop'.
© Apple Computer Inc. 1983-1999.
Screengrab, March 2001.

Figure 3.20 shows a crude example of 'aerial perspective', in which the foremost window of the Macintosh graphical user interface has higher contrast and saturation than the more distant windows. This may be an accidental outcome of employing weaker tones and colours to represent objects which are currently unavailable, but has a distinct depth-spatial result, albeit one with only two states: 'nearest' and 'all other distances'. The objects within the windows are unaffected by the algorithm.

Dunning points out (***op cit* p52-3**) that tonal value and figure-ground considerations have allowed artists to break the colour perspective 'rule' of foreground red with background blue in the past, for example using cyan as the foreground colour. To a certain extent the mere contrast of hues is enough to achieve the effect.

Perspective of blur

I noted that blur can be seen as one aspect of atmospheric perspective. However that phenomenon, the loss of definition with distance, is independent of the physiology of the eye, whereas Gibson's perspective of blur is the decreasing sharpness of the retinal image before and behind the plane of focus. The relationship of picture-making to this phenomenon will give clues to the relationship between space, time and realism.

The old view that the entire visual field is in focus at once is now modified in two

ways. First, the eyeball itself is known to change shape under muscular control in order to alter focussing distance (accommodation). Secondly, it is known that only that portion of the scene which is opposite the fovea is clearly resolved, and that it is through saccadic movements that this part of the eye is directed at different parts of the scene. So both in depth and across the scene, it is impossible for all parts of the scene to be equally resolved. However, by the nature of the eye's operation, that which is being attended to is generally in focus, and since we are not generally conscious of the eye's altering focus (and never of the saccadic movements) it could be argued that a representation which is in focus across its whole surface is true to our *experience*, and the history of painting is dominated by images which are entirely focussed both in depth and breadth. The experience represented is not that of an instant but is a summation of a period of observation.

We have therefore *two equally valid claims* to spatial realism. If differential focus is used in a picture, what will be the result? Will it be taken as true to reality or as a kind of 'encoding' which obtrudes into the relation between the scene and the viewer? The answer will be partly dependent on whether or not it is sufficiently marked to be noticed, but more importantly it depends on the culture of images to which the viewer is accustomed. As with the convergent verticals remarked on by Klee and Gregory, the perception of realism – of naturalness – will depend on how the viewer expects pictures to be.

Painters have represented differential focus with varied objectives. It seems certain that Rembrandt's self-portraits owe a part of their feeling of presence to his depiction of the differential focus of the planes in viewing the face, so that the eyes, the hypnotic subject for any painter observing his own face in a mirror, are in sharp focus, while the tip of the nose and the distant parts of the head are relatively defocussed. In this case what is being evoked is the visual subjectivity of concentrated study.

Figure 3.21.
Chardin: A lady taking tea (detail), 1735.
From Baxandall 1985 (cover picture)

Baxandall (**1985 p80**) claims that Chardin (Figure 3.21) uses selective sharpening and softening of edges in his paintings, in order to imitate the effect of the eye taking certain trajectories across the scene. The painter hopes to lead the viewer's eye through the painted image by increasing the sharpness of certain edges and points on a particular trajectory. Certainly Chardin makes frequent use of differential focus to

emphasise the central area of his works at the expense of a less well-resolved periphery. The paintings begin to represent the *process* of seeing, though themselves static.

Even within the rather tightly constrained practices of lens-based photography, the originator has the the freedom to exaggerate or to suppress blur. By choosing a wide lens aperture, the depth of field can be compressed, especially for near subjects; conversely a small aperture (given sufficient light or exposure-time) allows the depth of field to be greatly increased. The use of deep focus has the obvious informational advantage that it allows optimal clarity in all parts of the image regardless of distance.

In cinematography (and its synthetic equivalents in computer graphics) the depth of field can be altered dynamically, and the location of the focal plane can also be altered over time. There it has strong narrative potential, and helps indicate the important differences between author-controlled and user-controlled moving imagery (see the chapter on Film).

Relative upward location in the visual field

In everyday experience, one looks down at things that are near, and up at things that are far away. Experiments recounted by Gregory (**1977 p204-211**) showed that this effect tends to reassert itself even when artificial inversion of vision is imposed on subjects using mirrors or prisms.

Figure 3.22 **Van Eyck**: The Ghent Polyptych, c.1432 (detail).
From Faggin, Giorgio T 1986 *The Complete Paintings of the Van Eycks* Penguin Books, Harmondsworth UK, plate XXV

A striking example of elevation-as-distance (together with size, Gibson No.2) is provided by this panel (Figure 3.22) from a Van Eyck polyptych. Again the omission of some perspective cues is notable. Most obviously absent are any cues of diminishing tonal contrast, hue, saturation or detail, so that its spatiality represents a

compromise between the demands of geometric perspective and more schematic models like that of the stained glass Massacre (discussed above): in modern terms it might be said to be both a 'diagram' and a 'picture', since it affords viewing both as a planar configuration and as a natural scene.

Figure 3.23
Unnamed Chinese painting.
From **Hogben 1949 p192**

Figure 3.24
Aoki Mokubei (1767-1833):
Autumn Landscape
From a book on Japanese prints (in Japanese), Kodansha 1970, plate 134.

In these Japanese and Chinese scroll paintings (Figures 3.23 and 3.24), it is true that the upper reaches represent the further parts of the landscape (as we expect in Western art) but there are subsidiary depths which the viewer is invited to explore. The articulation of this depth is done using a variety of techniques, but the main principle is very un-Gibsonian: it relies on the identification of particular familiar elements. Only because we recognise paths, clearings, buildings and above all people are we able to begin tracing an often obscured route through the space and to identify the depths in the landscape. There are however in addition some obvious and familiar perspectival devices, especially the diminution of modelling, detail, saturation and contrast with distance. The Chinese painting uses the reversed colour perspective remarked on by Dunning: in the upper reaches the foreground is a cold blue-green while the mountains behind are a warm brown. In formal terms, this tends to counter any dominance by depth effects which might 'break' the planar qualities of the design, allowing it to be appreciated as a surface – as a painting – as well as a depiction of depth.

Hogben (**1949 p193**) uses the Chinese painting in Figure 3.23 to make a classic identification of realism with geometric perspective constructions, eliding (and favouring) the two, while treating other spatial cues as somehow beneath consideration. He remarks in a disparaging tone: 'The artist aloof from science, like the painter of this Chinese picture, is less concerned with objective reality than with the portrayal of atmosphere. His picture [...] reveals no incentive to apply the rules of perspective.' Hogben offers a teleological view of (linear) perspective:

> What was essentially new in the fifteenth century was the investigation of the laws of perspective and their application; and the investigation itself was the signal of renewed interest in experimental optics. In the Greco-Latin murals and vases of antiquity, in Gothic art and in painting of the Byzantine tradition, we meet with *various makeshifts to suggest depth*...
>
> Hogben 1949 p186-8 *emphasis added*

Similarly, Parsaye and Chignell suggest:

> The use of linear perspective to create 3D images was one of the achievements of the Renaissance. One can see just how revolutionary the use of perspective was by comparing pre- and post-perspective paintings of the same churches in Italy. It is like comparing pictures drawn by children and adult artists. In retrospect, one wonders how people could have accepted two-dimensional representations of people and objects.
>
> Parsaye and Chignell 1993 page 204

The problem here is that both commentators fail to recognise the purposes for which the various perspectives can be used or to appreciate the different effects that each produces. For example, I noted earlier that one of the difficulties of the powerful recession of traditional geometric perspective is the dominance of the centre: it is difficult to prevent the viewer's eye from succumbing to its pull. By contrast, these oriental paintings offer what might be termed an 'exploratory' form of perspective which offers opportunities to 'feel' one's way into the space. They tend towards a paced narrative mode, inviting the eye on a relatively slow journey around the space, rather than a rush to the horizon. This is not of course to suggest that these paintings are in their turn 'better,' simply that different spatial constructions afford different kinds of viewing. Figure 3.25 makes a very effective use of exactly the deep space avoided by the oriental examples, using strong linear perspective.

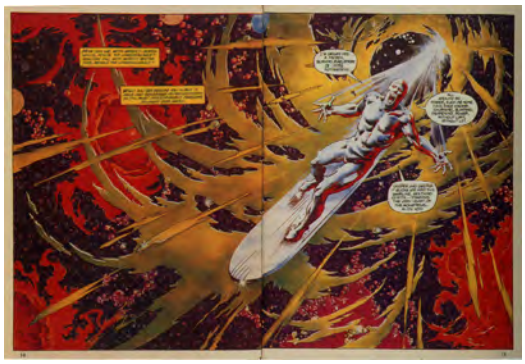


Figure 3.25
A double-page illustration by John Buscema for *Epic* 1980 Marvel Comics.
From **Sabin 1996** p154.

It is not suggested that the makers of each work have a conscious rationale for selecting particular forms of perspective. Clearly these choices are determined to a large extent by the surrounding culture. However in our own time, though never free from the limiting effects of our own culture, we are exposed to a huge variety of

different spatialities each affording different potentialities, and to a certain extent are free to choose the ones which suit the task in hand.

Sudden shift of texture density or spacing

A valley seen over a cliff is perceived as more distant partly because of the rapid change in texture frequency as the eye passes over the edge. The suddenness of the shift is an indicator of a possible contour and of differential distance at that contour. As with other cues, an alternative explanation for the stimulus might also be correct.

Figure 3.26.
Van Eyck: The madonna with Chancellor Rolin 1435 (detail).
From Faggin, Giorgio T 1986 *The Complete Paintings of the Van Eycks* Penguin Books, Harmondsworth UK, plate XLIII

In this painting by Van Eyck (Figure 3.26), one might expect to see both atmospheric and colour perspective used to differentiate the distance of the scene beyond the window from that of the interior. While both kinds of aerial perspective are used in this painting, they are reserved for the still more distant hills beyond the city. Within the detail reproduced here, the sudden change of texture scale at the contour between the near and the far is the most important cue, separating the scene into three planes: the figure, the battlements and the landscape. As in the Uccello *Battle* (with which it is roughly contemporary) the distance is spatially articulated as a painting in its own right, using a different dominant perspective cue from that of the rest of the painting. Yet to the general viewer of the painting, the fact that the depth cues are inconsistently applied is probably not noticeable.

Figure 3.27. **Van Eyck**: The Madonna with Chancellor Rolin 1435.
From Faggin, Giorgio T 1986 *The Complete Paintings of the Van Eycks* Penguin Books,
Harmondsworth UK, plate XLIII

In the painting as a whole (Figure 3.27), convergent linear perspective is used. This has the usual result that the background might seem to rush away from the scene, creating a 'hole' between the two actors in the scene, if they were not strongly tied together by the nearer form of the colonnade and if the strong tapestry-like detail of

the townscape did not tend to create a limiting 'backdrop'. In this respect, the flattening at the back of the space caused by the selective use of spatial cues helps to maintain the unity of the image.

The angel above the Madonna is an ambivalent object scaled on one reading according to distance and on another according to importance. As noted previously in Van Eyck's work, he combines some of the schematic qualities of earlier periods with a more optical form of depiction. This is of course to consider the image in anachronistic terms: there is no reason to suppose that it was conceived in this way. However, I suggest that the different kinds of viewing which it affords can reasonably be imputed to the original viewers.

I noted in the introduction the formative effects of the visual culture of which the picture-maker is a part. While personal innovation in many cultures is not necessarily a goal in itself, an important source of spatial development may be painterly 'difficulties' such as those in the Van Eyck, leading to *ad hoc* adjustments, inclusions and omissions. These reactive adjustments may originate new traditions, and may be rationalised into a consciously articulated system. Such theoretical systems may in turn lead to the development of practice. Gombrich suggests that...

> the history of art [...] may be described as the forging of master keys for opening the mysterious locks of our senses to which only nature herself originally held the key. [...] Like the burglar who tries to break a safe, the artist has no direct access to the inner mechanism. He can only feel his way with sensitive fingers, probing or adjusting his hook or wire when something gives way. Of course, once the door springs open, once the key is shaped, it is easy to repeat the performance.
>
> Gombrich 1977 p304

With modifications, this is a useful metaphor. It captures the *ad hoc* aspect of pictorial development as well as the importance of prior example and tradition. The danger of the metaphor lies in its implication that there is one solution towards which picture-makers are striving, the one correct combination which will open the lock. This would imply a single model of spatiality towards which all picturing is impelled, (perhaps that of photography?) and which we should presumably have reached by now. This would prevent explanation of the multiple forms of spatiality which I document here, which extend both across cultures and across the pictorial genres within cultures. The truer picture is of *making* solutions rather than *discovering* them.[18] Each depth cue is a key in its own right, and each unlocks a different door.

Shift in amount of double imagery

If one looks at a distant point, everything between the viewer and the point will be perceived as double. The closer to the viewer, the greater the doubling. This binocular phenomenon has caused controversy as to which is a truer model, that of two discrete retinal images being combined in the brain by a mass of computations which identify commonalities between the two images, or, operating in a more primitive way only on the invariant information provided by the two images (**Bruce, Green and**

[18] In relation to linear perspective in particular this view has been controversial. An extended discussion is offered in the next chapter.

Georgeson 1996 p264-5). I am not aware that this phenomenon has ever been used as a cue in its own right in the making of monocular pictures or designs, though it arises naturally with any binocular technology such as the stereoscope or head-mounted VR device.

Shift in the rate of motion

Whereas Cue No. 5 refers to optic flow considered as a gradual phenomenon, Gibson treats separately those distinct differences in the rate of motion which are cues for the contour, or edge, of an object. This differential movement of objects is that experienced when an observer moves sideways relative to a scene, the closest objects moving across the visual field faster than distant ones. In train travel, stations pass very fast, while distant villages move more slowly. This is the primary cue used in early scrolling computer games such as Sonic the Hedgehog which typically used three planes at differential rates.

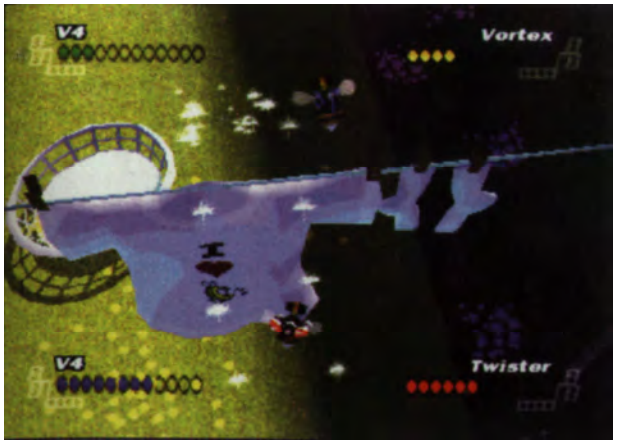


Figure 3.28.
MicroManiacs game for Sony Playstation, 2000, Codemasters.
The perspective is not convergent, but nevertheless very strong depth effects are achieved by the relative movements of the three layers of the model: in this scene, the players, the washing on the line and the lawn below.

Even recent games such as Micro Maniacs (Figure 3.28) forego linear perspectives in the interests of speed, but make extensive use of differential shifting of planes so that the user's player can be positioned convincingly at any height above a terrain. The fact that movement of the planes is essential for this effect is unproblematic in such racing games, whereas in a strategy game with extensive pauses the 'flatness' might be more apparent: instead the vertiginous effect of moving at a height above a terrain is accentuated with increase in speed and changes of direction, the essence of this kind of game. Again selected cues are chosen for the particular properties they offer.

Completeness or continuity of outline

If the outline of one object interrupts the outline of another, it will be seen as in front of the interrupted one. However, as Gibson puts it (**1950 p142**) 'a man knows that a near object can partially obscure a far object but his retina does not' and, seeking as usual the lowest-level implementation of perception, Gibson suggests that cues 9, 10 and 11 are the explanation for the ability to identify the edges of objects. However, he admits that the shape of objects is also significant. In Figure 3.29, it can be seen that **c** does not suggest occlusion, while **a**, **b** and **d** do. Such effects are explained by Hoffman in terms of a perceptual principle which favours generic or stable views (**Hoffman 1998 p25**) – those 'explanations' least fragile to changes of viewpoint. For

example **a** is explicable in terms of two short blocks abutting the sides of a longer one, but such a view would be obtained only from a unique position and would be lost on changing viewpoint. The view at **c** has the opposite properties: the most stable explanation is of two abutting angled objects, since many views onto such an arrangement would yield broadly similar images, while other putative configurations of objects would yield such a view only under unique conditions.

Figure 3.29. Perception of overlap is partly based on shape. Based on **Gibson 1950** p142, Figure 62

Figure 3.30.
Tomb of the Vizier Ramose at Thebes (1355-1350 BC) (detail).
From James, TGH 1992 *Egypt: the living past* British Museum Press, p192

The bas-relief in Figure 3.30 comes close to using only occlusion as spatial constructor: certainly size and linear perspective are absent. Texture gradients model the curvature of the dome of the heads. There are sudden shifts of texture density, but these are as strong whether they relate to changes across the surface of spatial entities or between entities. As with all bas-reliefs, the light and shade provided by shallow modelling stand for deeper space in the world depicted.[19] Completeness of outline is by far the strongest spatial cue here, and relies to a certain extent on world knowledge about the likely shapes of faces and bodies, though also on factors, documented by Hoffman (**1998 p33-4**) such as the occurrence of T-junctions where one contour disappears behind another, such as where the top of the further figure's head disappears behind that of the nearer.

A similar set of issues arises in relation to the perception of transparency (that is, the perception that a translucent object or filter overlays an object or scene) which is not covered by Gibson. Hoffman (**1998 p124-6**) makes clear that shape and position interact with tonal value in dictating whether a shape is seen as a filter, and a similar mix of influences affects the perception of layering and occlusion (***ibid* p47-51** and **p156-157**). Tufte devotes a whole chapter of Envisioning Information (**1990**) to

[19] The technique of bas-relief offers the *modelling* aspects of light and shade, but reliefs are normally too shallow to provide cast shadows, which in any case would tend to break the spatial effect rather than to reinforce it, since the cast shadows would fall in different places than they would in the world.

layering, which is often dependent on transparency (for example in Figures 3.31a-b). Further consideration of transparency in digital graphics is offered in the final chapter.

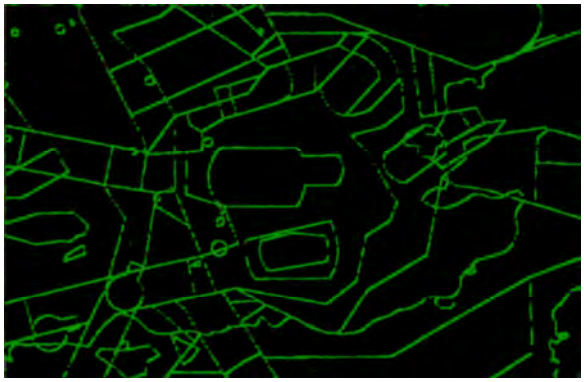


Figure 3.31a
A monochrome digital map for use in air traffic control. Note the degree of spatial ambiguity caused by lack of depth.
From Reynolds, Linda 1994 'Colour for Air Traffic Control Displays', *Displays* Vol.15 No.4, p217

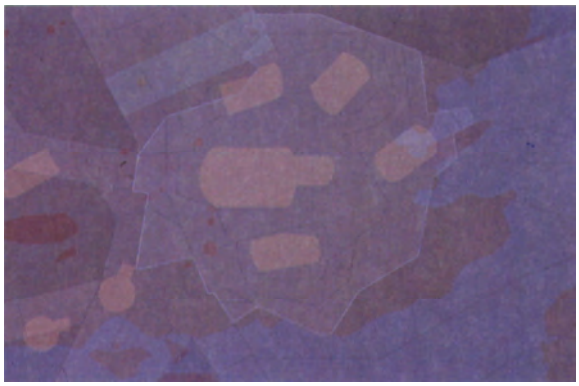


Figure 3.30b
A version of **a** using translucent colour. An impression of shallow depth is created and individual components become identifiable.
From Reynolds, Linda 1994 'Colour for Air Traffic Control Displays', *Displays* Vol.15 No.4, p217

Bertin commits the common elision of the concepts 'space' and 'geometry' which prevents him giving proper consideration to the kinds of spatial results of interaction between colour and shape to which Gibson and Tufte are sensitive. Bertin describes these 'retinal variables' as those which can be elevated above the plane, but he seems to mean this metaphorically: the original has 'variables rétiniennes que l'on peut élever en 3e dimension au-dessus du plan' (**Bertin 1973, p42**) but the later translation (made in close collaboration with Bertin) puts the word 'elevated' in quotation marks and omits specific mention of the third dimension (**Bertin 1983, p42**).[20] Perhaps he was right in feeling that it was confusing to use a spatial concept which might be interpreted metaphorically or literally depending on the reader, but it is symptomatic of his resistance to the kinds of depth cues which Gibson shows are so important to spatiality.

Transitions between light and shade

Gibson is somewhat confusing when he turns to transitions between light and shade, since he only includes this cue among his thirteen in the context of *abrupt* transitions at contours, whereas they should surely (as he seems to acknowledge elsewhere in the same work, **Gibson 1950**) also appear in the form of gradual transitions (as texture gradients do). It seems perhaps that his determination to distinguish retinal perception of the world from that of pictures (in which the use of shadow

[20] Bertin's insistence on separating the retinal from the spatial becomes eccentric when he discusses the difference between the use of lengths and areas to represent quantities, since it involves declaring length as spatial but area as not.

traditionally played an arguably disproportionate part) led him to play down the significance of this factor. Hoffman, despite a predilection for geometry which differentiates him strongly from Gibson, acknowledges the importance not only of modelling by shadow (**Hoffman 1998 p116-7**) but also of shadows which are cast (***op cit* p162-5**).

The case to which Gibson gives his attention in discussion of the thirteen cues is that of sudden transitions between light and shade at a contour, as for example depicted in Figure 3.32, where in particular the contrast between the light leaves and the dark masonry incites the observer to detect both an edge and a difference in depth. Such effects were perhaps less common before painting was influenced by the example of photography.

Figure 3.32.
Bowler: The doubt: 'Can these dry bones live?' 1855.
From Wood, Christopher 1981 *The Pre-Raphaelites* Weidenfeld and Nicholson, London, p69

Baxandall (**1995** *passim*) offers an exhaustive analysis of the role of shadow in depiction. The three basic forms of shadow are traditionally conceived as illustrated in Figures 3.32a-c.

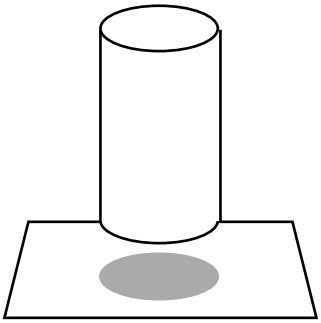


Figure 3.32a. Cast shadow

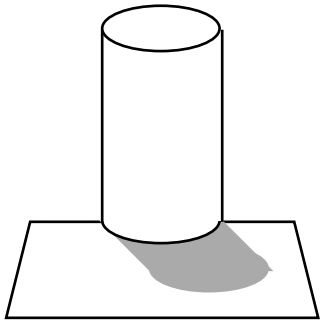


Figure 3.32b. Attached shadow

Figure 3.32a. Shading

A painter can choose exactly which shadows to depict, omit, accentuate or minimise. Computer graphics methods, in their simpler forms, have also made easy the presentation of one kind of shadow without others, since the shadows are calculated explicitly as individual projections of selected objects. Ironically in more advanced techniques such as ray-tracing and radiosity, such selectivity becomes more difficult, since shadows arise as a natural outcome of calculations based on the positions of all the objects and light sources. At the time of writing, it is still accepted practice in low-

cost virtual environment browsers and computer games to use shading without cast shadows, a liberty taken with spatial articulation for purely practical reasons of speed of computation.

Selective use of shadow within pictures is common. The objectives may be formal and compositional. Shadows may be deepened to increase drama, or lightened to prevent the obscuration of parts of the scene. It is important to note as with other cues, that shading and shadow techniques are effective means of constructing space even in the absence of some other spatial cues.

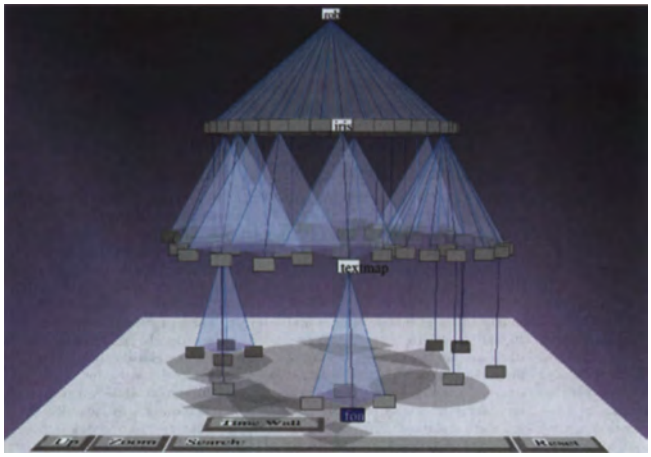


Figure 3.33
Robertson, Card and Mackinlay: Cone-tree visualisation of a directory hierarchy.
Card, Mackinlay and Shneiderman 1999 p525.

The synthetic shadows of Figure 3.33 have been designed to fulfil particular objectives. In particular, the fact that each shadow is computed independently of the others (so that where the shadows of two objects fall there is twice as much shadow and so forth) makes the shadows more useful than they would otherwise be for helping the user to locate parts of the cone-tree in space.[21]

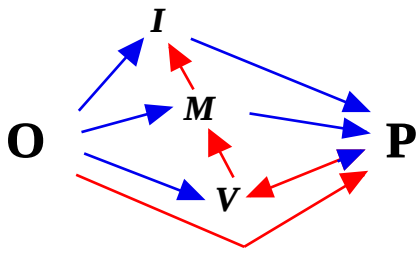
This chapter has been concerned with the range of pictorial methods available for suggesting depth, particularly the imitation of the sense of depth arising from looking at the real world. It centres on the relationship between the space of the representation and the space of the world represented. A number of points have begun to emerge which are pursued in the ensuing chapters.

The tendency has been noted in several authors to equate the evocation of depth with the techniques only of geometric perspective (linear and size perspectives). At its most extreme these approaches treat images which use such techniques as uniquely

[21] With a single light source and a set of translucent objects, shadows will behave as in this system. However that is not the main point: the shadow-casting algorithm has been chosen to capture those aspects of the scene which best suit the objectives.

correct; from here it is a small step to equate *correct* with *desirable*, regardless of the objectives of the image. By contrast, I have argued that selecting particular techniques allows particular objectives to be served in the resulting picture. In the work of Van Eyck it allowed a picture to be read schematically as well as pictorially. In the oriental landscapes it allowed the space to be explored in a longer and more involved process than if depth had been more forcibly depicted. Suppression of aerial perspective permitted distant objects to be more clearly discerned, and also in several cases allowed the picture to be more readily seen as both scene and surface, perhaps to make its formal qualities more apparent. In the de Chirico it allowed the 'game' of spatial depiction to be made a part of the subject. The pragmatic explanation – that the selection of depth cues serves the objectives of the image – seems better able to rationalise the kinds of figurative images which have been made.

One question which has arisen is of where in the ***OIMVP*** model diagrammed earlier decisions about depth cues are effected? Many occur in the process of picturing, ***P***, when particular kinds of rendering and mark-making are chosen. However many also arise in ***V***, the view selected, (as for example when a particular view is chosen on the model to afford some depth cues rather than others) and even in ***M***, the model, as when Uccello contrives particular subject matter, and positions for that subject matter, so as to provide material from which to construct orthogonals. Often there is a diegetic excuse for a pictorial tactic. For example, in the Watteau (Figure 3.18) the foremost figure 'happens' to be dressed in red, reinforcing aerial perspective. In terms of the earlier diagram, such manipulation of the depicted matter in order to help produce a particular picture might be indicated by various reversed arrows to indicate how pictorial decisions feed back to apparently 'prior' stages. This symbiosis of model, view and picture will be seen later to be fundamental to the spatiality of film and is discussed more fully there.



From the Renaissance onwards mainstream perspectival painting seems to move towards making the image *apparently* consistent with natural vision, even when it is nothing of the sort (and despite the fact that an increasing range of other kinds of images is also made). This difference between graphic images which are designed to be taken as unencoded representations and those where the pictorial intervention is more overt will also turn out to be important in the discussion of other media. In the next chapter I will discuss how images which are clearly on analysis not imitations of the optical impression of a scene can be taken to be realistic.

I have begun to show that the selection of depth cues has consequences which are

both informational and affective. Sometimes the two are in accord, as when the 'flattening' of a Chinese landscape provides information about distant parts of a scene which might not actually be discernible in reality, at the same time creating a particular relation to the viewer. Sometimes a particular approach, say binocularity, offers little informational advantage but a substantial one in terms of affect. Of course as already indicated in the Introduction, the difference between information and affect is not a strict one: the use of blur to capture the differential focusing across a scene may be conceived as conveying information about a visual experience as well as in part recreating that experience in way which alters the viewer's relationship to the image. However one can probably safely say that it it does not provide more information about the depicted world; indeed it provides less, since it suppresses detail in parts of the image. This question of whether any given picture offers information about the world or about the process of seeing the world, forms an important part of the next chapter. There the discussion is informed by the difficulties raised in this chapter by the rival claims to truth which have been revealed in attempting to apply the depth cues of natural vision to pictures. These have included possible mismatches of 'logical' and 'psychological' truth (to borrow Klee's terms for the time being) in relation to size perspective, verticality, focus in depth and focus across the visual field. Already it seems possible to detect important inconsistencies in the apparently logical space of the archetypal post-Renaissance picture (for example that world verticals are to be made vertical in the picture, but horizontals are not so treated). However before this can be asserted with confidence it will be necessary to resolve some traditional controversies over the claims of certain projection systems to be considered correct.

The utility of that correctness, if established, also requires discussion. I have already noted that it seems more important for the internal spaces of a picture to construct an apparently coherent scene than for the image to match the visual stimulus available if the user were present at the scene. It has become apparent that different kinds of pictures may have their own claims to be considered 'realistic'. However, it also seems that the ability of any representation to fulfil the objectives intended for it is not always increased by the use of greater realism. In the next chapter I will show that the mismatch of pictures to optical truth (in so far as it can be established) often enhances their expressivity. This will again emphasise the pragmatic approach characterised by Gombrich as *making rather than matching*.

4 Spatiality and realism

1 Introduction

In any discussion of how the planar space of pictures corresponds to the volumetric space of the world, it is often assumed that some kinds of picture have a special correspondence to how the world actually looks. Some analysis of the concept of *realism* is therefore unavoidable. Unfortunately the term is habitually used without proper definition, based on careless assumptions about matching pictures to scenes.

The argument hinges on two issues (1) whether it is possible to match optical truth (which must be defined) and (2) the objectives in attempting or not attempting to do so.

In relation to the first, the argument put forward is opposed to the relativist position that no particular kinds of pictures have a better claim than another to be considered correct. In particular, by looking specifically at different spatial projection systems, I show that one system does have a superior claim to match the scene. In the process I highlight some misunderstandings concerning the planar projection of scenes which have made the relativist position seem more convincing than it otherwise might.

The geometry of spatial projection is not however the only means by which depth is depicted, as indicated in the previous chapter. While a particular projection system may have a superior claim to correctness, this does not mean that every aspect of some spatial depiction can be indisputably claimed as realistic. The difficulty of achieving an unequivocal definition of optical truth discussed in the previous chapter becomes important here. If there were a kind of image quintessentially realist because it was the visual equivalent of looking at a scene, such an image would necessarily operate independently of any sort of codes or conventions of depiction and picture-making would not be the pragmatic activity I have claimed. If on the other hand it can be demonstrated that no picture is a uniquely realist representation because all pictures must adopt some particular conception of realism, then pictures are essentially *designed* artefacts which will often require shared prior knowledge on the part of maker and user and which are created within a purposive context. As a result, by the end of this chapter, the simple concept of realism will have been replaced by that of *kinds of realism*, adopted for the purposes they serve.

The second objective is to consider what the *uses* of various spatial realisms are; particularly, the reasons why picture-makers may depart from any goal of straightforward imitation in the design of pictorial information. If some pictures come very close to being 'right' as representations of scenes, the question remains of what purposes other kinds of pictures serve. I shall propose that forms of realism that seem to 'work' (they *seem* real) are as important as those which might be said to be 'true'. I contrast attempts to straightforwardly match scenes with another approach to spatial realism which I define. This alternative approach to realism particularly

reflects the role of *time* (considered in a variety of ways) in static pictures.[1] It is argued that departure from a simple notion of realism is potentially more expressive both in informational and affective terms.

2 Kinds of realism

In the attempt to discern what visual realism may be, I do not want to discredit the term *realism* as a useful shorthand for some approaches to image-making. But the term has dangers when it is used without being defined and embodies hidden, unconsidered assumptions. In both the analysis and making of pictures I suggest it is better to replace the vague general idea of realism with a model of multiple realisms, some of which are closely tied to the different depth cues discussed previously.

Some of these might be:

1 depicted 'tactile' qualities such as the bloom on fruit, where the realism is such that spectators feel they could pick up the fruit and eat it (**Bryson 1990 p28**); this is related to spatial characteristics, but is not itself essentially spatial.

2 the sense of 'objectness', such as a painted curtain having the concrete presence of the real thing (as in Pliny's account of the competition between Parrhasios and Zeuxis in Natural History XXXV p65 quoted ***op cit* p30**), or the trompe l'oeil of Cornelius Gijsbrechts (***op cit* p142-4**); this is a predominantly spatial phenomenon; it often carries a notion of the 'weight' and groundedness of the depicted object.

3 the sense of solidity, produced largely by modelling with light and shade; this can be considered for an individual object, to a certain extent independent of context (***op cit* p66**); it is a predominantly spatial phenomenon.

4 a sense of depth produced by an illusion of projection, achieved principally through the use of cast shadows and of the positional interrelation between objects (***op cit* p67**); this is a predominantly spatial phenomenon.

5 'occupiability', extending the space occupied by the spectator (***op cit* p34**), often through the medium of geometric perspectival construction, but also achievable as Bryson notes (***op cit* p42-3**) using colour, tone and scale without linear perspective; a predominantly spatial phenomenon.

To this list can be added other impressions of the real, some of which may be assisted by particular kinds of spatial representation, but in which there is no simple, direct implication for spatiality:

6 a sense of movement, as when a static image is able to suggest moving water or wind in the trees.

7 a sense of animation, of coexisting with a living being ('the eyes follow you round the room', 'she looks as if she's breathing').

8 a psychological engagement as though with another mind.

9 general recognition, as in 'that's a...'.

10 specific recognition, as in 'that is definitely an example of *x*' when an image is so accurate that one can identify, say, a particular variety of tulip (**Bryson 1990 p106**).[2]

[1] It might seem that this would make the findings inapplicable to representations in which *actual* time plays a part – film, television and interactive media. However, rather surprisingly, this turns out not to be the case, an issue pursued in the remaining chapters.

[2] There are really more than two levels of recognition, perhaps an indefinite number from the broadest notion of what something is to its exact identification as some highly particular thing, but two levels suffice for the discussion here. There are also degrees of certainty of identification.

The two kinds of recognition 9 and 10 are based on different visual attributes, depending on what is depicted. Illusion of spatial depth may or may not be significant. For example, in the specific identification of the beer bottles as bottles of Bass in Manet's *Bar at the Folies-Bergère* (Figure 4.01), the identification is based on a logotype: the depth illusion is irrelevant to this recognition. However, the disputed general recognition of the scene behind the barmaid as a reflection or an extension of the room is clearly primarily a spatial one. I indicated earlier that recognition is itself an important secondary cause of depth perception, so a circular relationship is implied.

Figure 4.01.
Edouard Manet: A Bar at the Folies-Bergère, 1882
From Richardson, John 1982 *Manet* Phaidon Press, Oxford, plate 46.

General recognition forms the principal basis of yet another kind of realism:

11 being 'like everyday life'; this is the realism of Courbet (**Hanson 1979 p33-34**) or the Soviet realists of the 1930s (**Taylor 1987 p135-6**) where there is a sense that what is depicted is in a broad but rather particular sense 'like real life'. Choice of subject matter is decisive here, but the manner of representation has often in practice been regarded as material to the overall effect.

12 the final realism is that of causal coherence, which allows the observer to predict likely outcomes. This may have a strongly spatial element, as for example, in Gombrich's example (**1973 p228**) of a picture of a ship, the movement of which is anticipated by the observer. Gombrich remarks (***op cit* p218**) that 'the truth we seek with our senses is not the static and eternal truth that interested Plato, but the correct assessment of the developing situation with which we interact.'[3]

Relations between realisms

These various realisms may be used selectively. For example, the realism of solidity may be achieved through modelling alone without the need for cast shadows. Conversely a coherent, occupiable space can be pictured using cast shadows but where modelling is absent. However, such different spatial realisms are often also used together. The case of traditional Japanese printmaking, in which modelling is absent, is one where cast shadows are absent too. Masaccio, though he used local modelling to create solidity in the absence of a coherent shade model (**Dunning 1991 p57-68**) as in the Expulsion from Paradise (Figure 4.02), is also celebrated for 'sculpting' with light, in which cast shadows and relative tonal values articulate an

[3] Prediction implies an important role for time in relation to pictures. In this chapter this concept is largely confined to the incorporation of time into pictures, while in the next the complementary issue of pictures in time is explored.

'occupiable' space, as in The Tribute Money (Figure 4.03). The absence of local modelling here is difficult to imagine.

Figure 4.02 (left)
Masaccio: The Expulsion from the Garden of Eden 1427, Brancacci Chapel, Florence. From **Dunning 1991** p66**.**

Figure 4.03 (below)
Masaccio: The Tribute Money 1427, Brancacci Chapel Florence. From **Dunning 1991** p66**.**

Peter Greenaway notes the range of realisms which operate simultaneously in Holman Hunt's *The Hireling Shepherd* of 1851-2 and it seems clear that they do not operate independently of one another '...the sleeping sheep is heavy, the green apples are bitter, the grass in the ditch is wet, the woman's feet are palpable. With no trouble at all you can walk about the painting like you can walk about a landscape – there is enough evidence to name all the plants' (**Pascoe 1997 p34)**. Here we have the sense of weight associated with 'objectness' and modelled solidity (the sheep), the inference of non-visual qualities such as taste (bitter apples), feel (wet grass) and impressibility (palpable feet), occupiability (the landscape that can be walked in) and specific recognition (evidence enough to name the plants). This seems to raise the possibility that there is a kind of realism which 'has everything' – a super-realism which subsumes all the realisms so far enumerated. However, it must be pointed out that the list is still incomplete in an important respect. The realisms listed so far have principally to do with a single static moment, so that any element of time-awareness – such as the sense of movement in depicted natural phenomena – evokes the passing of time in the observed world. None of these realisms captures the fact, noted often in the discussion of the Gibsonian cues, that *perception itself* takes place in time. This is crucial in delimiting the extent to which any kind of pictorial visual realism can be achieved. The relationship of time to pictures is discussed at some length below. In addition, aspects of realism which seem far removed from any discussion of spatiality, such as psychological realism, will turn out to be of direct significance to spatial depiction when film is discussed in the next chapter.

Using small subsets of realism

There are many examples where only a small subset of realism is invoked, for example in graphic artefacts such as the computer interface. In Figure 4.04, the colour circle and the colour chips at top right have drop-shadows producing a sense of a

'real' object but with no corresponding sense of a coherent space which the object occupies – some objects have shadows while others have none, or have shadows of a different kind. Despite these inconsistencies, a sense of palpability or objectness is conveyed. The spatial qualities of the virtual slider below the colour circle are of a more conventional kind, depicting a relatively 'believable' object in that one could envisage such an object existing outside the digital surface. The graduated colour strip below it offers a probably unintentional and unwanted depth illusion of its own – a reminder that depth cues assert themselves unbidden even when no depiction is intended. In an informational sense, the cast shadows provide nothing at all, but in terms of affect they help to make the display tangible, more 'real'.

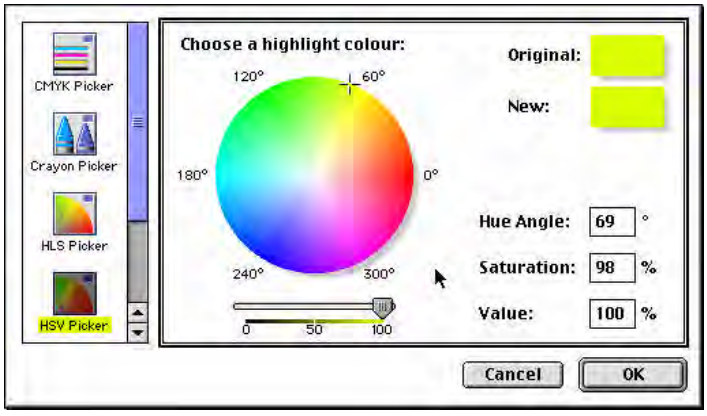


Figure 4.04.
Colour picker interface object from Apple Macintosh Operating System 8.1 ©Apple Computer 1983-1999

The same shadow motif seen in the Apple interface is frequently encountered in the work of the publisher Dorling Kindersley (Figure 4.05a).

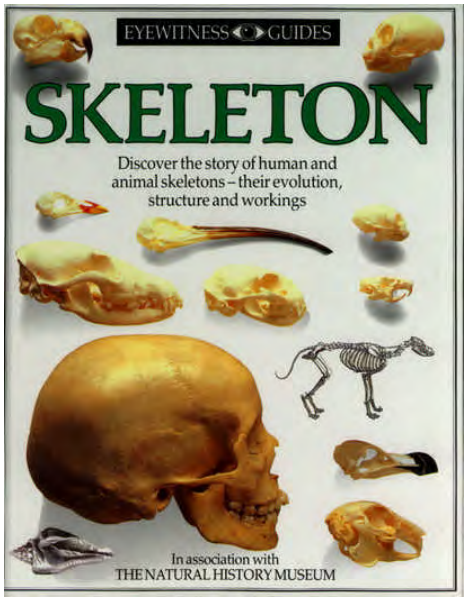


Figure 4.05a.
Book cover 'Eyewitness Guides: Skeleton' Dorling Kindersley, London 1988, author Steve Parker.

Close examination shows that many of the shadows are digitally originated or edited. They do not make coherent sense (Figure 4.05b) but are used to enhance the realism of objectness, perhaps with a view to diminishing the remoteness which traditional 'bookishness' otherwise imposes on book illustration. Many kinds of realism are used for these affective rather than strictly information purposes.

Figure 4.05b.
Detail of R.05a.
Does the engraving of the shell float above the surface on which the skull casts its shadow? Why does the engraving not cast a shadow of its own? Why does the dog skeleton in R.05a cast no shadow?
Spatial articulation has been contrived on an *ad hoc* basis to make the book cover 'work' without reference to an imaginable coherent spatial referent.

Such contrived spatialities are not new. In a drawing of 1526 by Albrecht Dürer (Figure 4.06), much of the internal modelling of the figure is imaginable in a real scene, but the halo of yellow is entirely synthetic, being used to make the figure 'come off the page'.

Figure 4.06.
Albrecht Dürer, Study of a man walking in profile to the left, 1526
From **South Bank Centre** 1997 The Quick and the Dead: artists and anatomy (book to accompany National Touring Exhibition 1997-98) South Bank Centre / Arts Council, London, p80

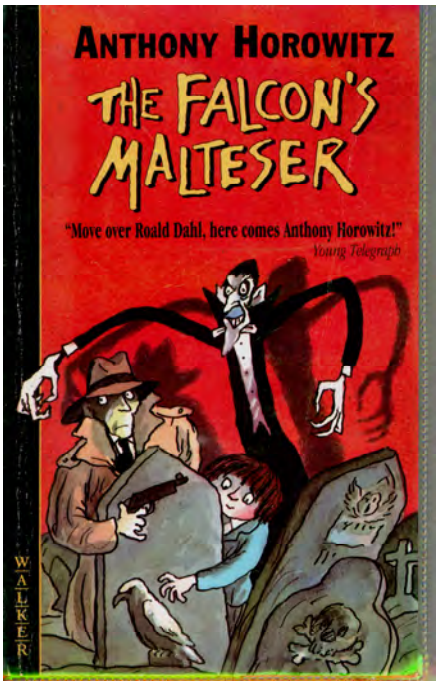


Figure 4.07.
Cover by Tony Ross for The Falcon's Malteser by Anthony Horowitz, published by Walker Books 1986

It is an example of *illicit mark-making*, constructing space rather than depicting it in any straightforward sense, which I discuss later. A similar space, using some aspects of realism but rejecting others, is shown in Figure 4.07 – note the cast shadows.

Figure 4.08.
Weeping skeleton (Skeleton in landscape) attributed to Stephen Calcar, 1543 from Vesalius' Fabrica
From **South Bank Centre** 1997 The Quick and the Dead: artists and anatomy (book to accompany National Touring Exhibition 1997-98) South Bank Centre / Arts Council, London, p28

In this illustration from Vesalius (Figure 4.08) the sense of the real is created differently. A coherent, occupiable space is created by borrowing the conventions of landscape pictures. It was presumably done partly (and by many others than Vesalius) as the most effective way, in the context of the time, of making things 'feel real'. The scenery was imported with the purpose of making space – and that was done to support the sense of the real.

To achieve any particular form of spatial realism, then, differing cues may be called into play. In Cotàn's *Quince, Cabbage, Melon and Cucumber* (Figure 4.09), tactile realism is largely dependent on the texture and shade gradations. The sense of 'objectness' relies on the same cues, but also on cast shadows. Solidity depends principally on shade gradation. Projection is most dependent on size comparisons, completeness of outline and tone contrasts at contours. Occupiability is particularly promoted by linear perspective.

Figure 4.09.
Cotàn: Quince, Cabbage, Melon and Cucumber, 1561-1627.
From Bryson 1990 p67.

The cues do not *produce* the various kinds of realism: realisms are arrived at through the interaction of a mutually supporting set of cues with the content (that is, the

subject matter) and the context, the expectations and prior knowledge of the viewer and the picture-maker. While some of the realisms enumerated are mutually supportive some are found to be incompatible. Porter and Susman (**2000**) note the tensions in the making of Pixar's computer graphics films between subservience to visual realism and believability of character (*animation* and *psychological engagement* in my tentative list of realisms). In this case the need to be selective of realisms and their supporting spatiality is acknowledged. However, picture-makers often narrowly select the realisms to which they subscribe, while at the same time claiming an ill-defined general realism for their work: examples from digital picture-making are given next.

3 Contingent definitions of realism in digital pictures

In the many computer graphics papers which deal with 'realism', explicit definitions are rarely offered. Realism is taken as an uncontentious given; yet from each paper emerges a different implication of where realism lies. In general, these implicit definitions emerge from the purposes for which the image will be used, though in a few cases the image is seen as an image *per se* with no declared purpose.

A recurrent theme in the realism of digital imagery is, understandably, that of computational economy. In some cases this has no implication for realism since it merely involves eliminating the calculation of attributes outside the human visual range (**Greenberg 1999 p51-2**). In other cases sub-optimal realism is adopted when, though the shortfall will be detected, this is considered unimportant. For example Heidrich *et al.* (**1999**) comment: 'In flight simulators and other outdoor sceneries, fog can significantly contribute to the realism of a scene... [so a formula is used] to *fake* emission and scattering effects.' (***op cit* p130** *emphasis added*).[4]

Some means of balancing visual realism and computational economy can be considered as offering *degrees* of realism on which a numerical value can be placed, for example polygon count in models or levels of recursion in ray-tracing. However others cannot be quantified in this way and it becomes clear that individual visual realisms are being selected.

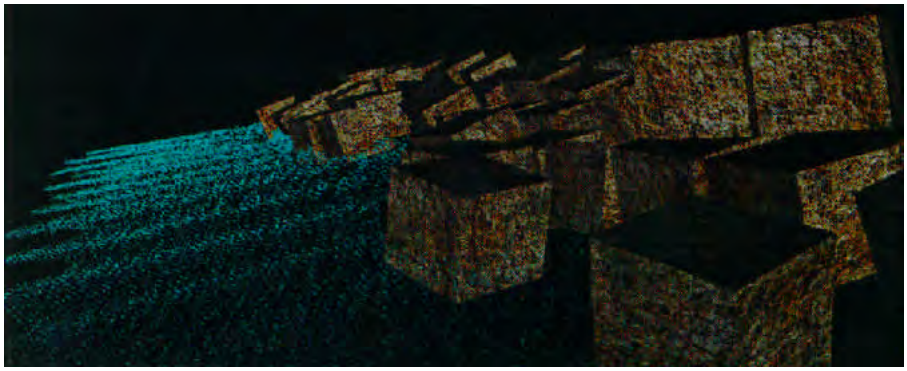


Figure 4.10. Computer modelled harbour blocks, representing Ijmuiden harbour, Holland.
From **Chapman *et al.* 1998 p572.**

[4] In relation to the graphics of computer games, popular journalism has a strongly pragmatic attitude to visual realism, even though realism is something of a talisman in this community. 'A big fat, hairy deal has been made out of Tekken Tag and the way the grass actually moves on one of the levels. So what? What does this add to the gameplay? Nothing.' (**Ellis 2000 p36**).

Chapman *et al.* (**1998**) describe an application of seabed visualisation for documentary purposes in capturing the movements of harbour-wall blocks (Figure 4.10).

When the authors report that 'The blocks are texture mapped with pseudo concrete and a lighting model is added to the underwater environment to improve the *realism* of the harbour model.' (***op cit* p480** *emphasis added*) there is no indication of what this realism is for. It is perhaps to make the simulation more 'believable' in some generalised way, though this is not mentioned in the text and the project seems to be intended for technical experts whom one might expect to be content with the unadorned geometry of the blocks. The accurate representation of the *geometry* of the blocks is crucial to the usefulness of this project, but one wonders what the criteria are for the selection of these other particular aspects of visual realism.

Figure 4.11.
Original caption reads: 'High and low oblique, computer-generated perspective views of the area around Irish Canyon, CO. Source data were composed of 20m SPOT imagery and polygonalized 10m terrain elevation data that were vertically exaggerated by a factor of 5'.
From **Weinhaus and Devarajan 1997 p354**

By contrast, the work of Weinhaus and Devarajan (**1997**) on texture-mapping 3-D synthetic models of real world scenes offers a clearer declaration of the aims of their particular style of realism, which effectively defines the criteria for its implementation: 'Today, more demands are being placed upon visual simulators to achieve yet a higher level of *realism*. In particular, mission planning and rehearsal systems are now striving for *truly faithful* representations so that ground troops can become intimately familiar with important regions of the world' (***op cit* p326** *emphasis added*). This amounts to a functional definition of one kind of realism, in which the sense of particular place, a feeling for how landmarks in the scene relate to each other and to the situated observer, is paramount. Intriguingly, the paper is illustrated with a landscape image in which the heights are vertically exaggerated by a factor of 5, even though other aspects of the image are broadly photorealistic (Figure 4.11). This is not justified in the main text.[5] This distortion suggests again an unacknowledged interpretation of spatial realism, possibly with a view to resolving a perceptual problem of size constancy, discussed later in this chapter.

[5] 3D geophysical relief maps typically use 10-fold exaggeration of heights, but there are more complex examples. A relief map of France by the Institut Geographique National uses a planimetric scale of 1:1,200,000 with a non-linear altimetric scale from from 1:80,000 to 1:160,000 (that is, between 15 and 7.5 times the planimetric scale). This presumably ensures that low-altitude differences are made visible while keeping the mountains manageably low.

Some authors offer a largely functional objective for the inclusion of particular spatial cues, in which the mimetic motivation is minimised. Schöffel (**1997**) and Soler and Sillion (**1998**) both justify their work on shadows not on the grounds that shadows are observed in the world and are therefore included in the image by right of realism, but because of the information they impart about the relative positions of objects, light sources and viewer. This is clearly a case of making space rather than matching vision.

Though they do not explicitly say so, Granieri, Crabtree and Badler (**1995**) seem to define realism primarily in terms of one particular form of believability. In particular they deprecate techniques which lead to high optical realism at the expense of believability in human motion behaviours: 'the visual effect of even the most perfectly animated figure is significantly reduced once the viewer recognises that its movements are exactly the same each and every time it does something' (***op cit* p238**). They prefer to store less information while allowing motions to be modified on the fly to match the context in which they are replayed. They are optimistic about the wider application of their techniques. 'Although today primarily driven by military requirements, the general technologies for projecting real humans into, and representing simulated humans within, virtual environments should be widely applicable in industry, entertainment, and commerce in the near future' (***op cit* p223**). However it is clear that a form of realism desirable for one application is not necessarily so in another. The question whether the particular realism which these authors have selected pragmatically to serve the needs of their product can be generalised to other different genres highlights the contingent nature of 'realism'.

Diefenbach and Badler (**1997 p60**) offer techniques for producing 'realistic images'. They use the term 'realism' freely without defining it. They are making images for their own sake – using realism as the yardstick of algorithmic ingenuity – rather than with any particular functional purpose in mind. They must be assumed to be equating photorealistic images with the perception of reality, though they also point out that some 'cheating' is possible, for example: 'Often, however, shadows are desired to simply provide visual cues and *some level of realism*, and are not required to be completely accurate.' (***op cit* p68** *emphasis added*). Like almost all workers in the field, they acknowledge that realism can be tempered by economy (***op cit* p69**).[6]

A recent issue of a professional magazine for computing specialists[7] featuring realism in digital images is beset with confusions as to whether the realism to which some computer graphics aspire is resemblance to pictures (such as photographs) or resemblance to the experience of looking at scenes. Just why it is dangerous to equate photographic imagery with realism will become clear in the later discussion.

---

[6] It is ironic that while the pioneers of photography took pains to explain to their public that 'groups of figures take no longer to obtain than single figures would require, since the camera depicts them all at once, however numerous they may be' (from 'The Pencil of Nature' by WH Fox Talbot 1844-46, in **Frizot 1998 p62**), computer graphics workers must constantly remind the reader that greater (photo)realism requires more complex and more subtle, and therefore more expensive, computing.

[7] Communications of the ACM, August 1999, Vol. 42 No.8

The aim, as introduced by Rosenbloom (**1999 p30**), is 'to create images believable to the observer – ultimately emulating and predicting reality.' When he says that 'the framework for realistic image synthesis developed at Cornell University's Program of Computer Graphics maintains fidelity between rendered images and their *physical counterparts*.' (**Rosenbloom 1999 p30** *emphasis added*), it seems however that these physical counterparts are photographs.

At first sight Greenberg's article on work at Cornell is more clearly defined, referring to a goal of making 'synthetic *images* visually and measurably indistinguishable from real-world *images*' (**Greenberg 1999 p45** *emphasis added*) but shortly afterwards a picture caption suggests: 'A scene observer receives the scene radiances and has a particular visual experience. Because we want the display observer to have the same visual experience, the displayed image is a perceptual match to the *scene*.' (***op cit* p51**). Now it is the matching of scenes and not the matching of pictures which is under discussion. The expression 'visual experience', is left undefined.

In most of the literature, realism is assumed to be self-defining. However, Usoh *et al.* (**1999**) not only declare their criteria in terms of the response of users but assess their success against these definitions. Their goal is the development of 'natural and effective virtual surrogates for user interactions with physical spaces and objects.' (***op cit* p359**). They use a number of forms of interaction with the space (theirs is a virtual space containing an alarming drop or pit) to achieve a 'strikingly compelling virtual experience' (***op cit* p363**), and they assess the user's reactions as evidenced by:

- the subject's awareness of background sounds
- subjects' reports of similarity to really looking into a pit
- subjects' sensation of vertigo and willingness to walk over the pit
- whether or not subjects actually traverse or circumnavigate the virtual pit.

This is, very unusually, a properly defined sense of realism based principally on the notion of 'believability' – in other words a definition based as much on the relation between representation and user as on that between representation and referent.[8] The subjectivity of the experience engendered is acknowledged in their finding that the sense of realism declined in proportion to the subjects' prior experience of computer gaming (***ibid***). Such recognition that the sense of realism is subject to change over time and to context is highly unusual.

Dykes, Moore and Fairbairn (**1999**) take a different approach. For them, deviation from visual realism is judged to be not only expedient for technical reasons but also in many cases *desirable*. Their cartographic tradition perhaps causes them to consider more clearly the role of audience and objectives in determining the uses of realism; even so, they perhaps regard realism as something that one has more or less of, rather than being itself defined by the objectives. For them the making of a representation includes 'selection, classification, simplification, exaggeration and symbolisation...' (***op cit* p99**). Above all, they say 'cartographic visualisation is not just about authentic replication' (***op cit* p101**), though it is unclear whether or not they believe that

[8] I explain later why the realism of a representation must be defined in relation to the observer as well as to the scene.

authentic replication is possible. They propose that 'the argument for replicating "The Real World" in the map, by producing authentic versions of reality from which individuals can extract pertinent information, is a strong one for use in a variety of tasks. The cartography in such maps involves the selection of information and provision of forms of interacting with it that are *suitable for the intended use*' (***op cit* p103** *emphasis added*).

Dykes, Moore and Fairbairn's view represents a new strand of argument – explicit reasons for *rejecting* aspects of realism even when it is achievable. This brief survey of the computer graphics literature seems to suggest three different motives for holding back from realism in some respect: economy of computation; distortion or exaggeration (such as the differential scaling of heights in a landscape) for reasons of expressivity; selectivity for reasons of expressivity, adopting selected realisms suited to an intended use. These may combine, so that for example expressivity and economy together cause the selection of some particular set of realisms in preference to others.

4 Towards a definition of visual realism

Returning to the particular question of the relation between the space of the world and the space of pictures, is there some kind of realism which has a special claim to be considered truly 'realistic' in the sense that it is a perfect substitute for natural vision? If this were so, then this kind of realism *could* be defined solely in terms of the referent (which it would perfectly match) without needing to include any consideration of the response.

One of the obstacles highlighted by Gibson – and now generally acknowledged – is that natural vision *depends on* movement. Many of the ways in which spatial depiction in paintings departs from that of photography will turn out to be related to this problem, as picture-makers grapple with the problem of representing what is essentially a time-based process in a static artefact. Because natural vision incorporates immersion in, and dynamic response to, the surrounding scene, the simple answer to the question 'is there a kind of picture which matches vision?' will be No. But can a picture (whether painting, photograph or digital graphics) be designed to match the appearance of the world in other respects, for example in terms of spatial geometry? The answer that will be proposed here is that (1) there is a kind of spatial realism which, under highly controlled conditions, can be regarded as equivalent to natural vision, but that (2) the many shortcomings of this kind of illusionistic picture-making are sufficient to make its special status almost inconsequential when choices must be made between kinds of depiction. This will help to explain why very few pictures conform to this particular kind of realism, which will be referred to as PI-realism (referring to *pictorial illusion*). The other – really a family of related approaches – will be referred to as VE-realism (referring to *visual experience*). In particular I will suggest the VE-realist images are often more *expressive* than their PI-realist counterparts.

5 Defining PI-realism

Gombrich suggests (**1973 p194 *passim***) that it would be intolerable to limit the definition of illusion in painting to the narrow case in which the observer actually mistakes the picture on the gallery wall for a window opening on a real scene, but this I suggest is a good starting place, a kind of *Turing test* of realism.[9]

Though Gibson tried subsequently to distance himself from it after developing his ideas concerning motion and the optic flow (**Gibson 1979 p227-229**), his conception of the 'faithful picture' is a useful one:

> A delimited surface is so processed that it yields a sheaf of light-rays to a given point which is the same as would be the sheaf of rays from the original scene to a given point.
>
> **Gibson 1954**

Gibson sought to disown this definition when he recognised the difference between what I have named PI-realism and VE-realism, but as a definition of PI-realism *per se* it still stands. A working definition of PI-realism then is of an image which:

- is geometrically automorphic: all parts of the image on the retina occur in the same places for the picture as they would do for the corresponding scene[10]
- stimulates the same sensations of colour and tone at every point as would be stimulated by the scene
- does not offer any stimuli *not* offered by the scene.

The spatial implications of such a form of realism will be dealt with here principally in terms of the controversy over rival projection systems, since it has been disputed whether spatial projection of the three-dimensional world to the two dimensions of the plane is a matter of correctness or of convention.

Alpers claims that the seventeenth-century Dutch did not distinguish seeing from picture-making (**1983 p27**). A descendent of this identification of certain kinds of pictures with natural vision is the notion that photography shows how something 'really looks' and I have shown that it is often an aspiration of computer graphics. This enthusiasm for perfect mapping of spatial reality to a surface is taken by some to be an extreme form of the 'hegemony of vision' (**Levin 1993**) introduced in Chapter 2. This hegemony is characterised by theorists as a belief that visual realism is the standard of realism in general, that how things look is a supremely important aspect of how they are, a point of view said to be unique to Western culture. Lalvani (**1996 p1**) co-opts Heidegger, Derrida and Nietzsche in alleging this hegemony and cites the opinions of Plato, Aquinas and Descartes as evidence. Vision is said to be prioritised as the mode of perception which, 'more than any other, perceives things in the world as objects that are clearly there, present, and available for us to study and use.' (**Houlgate 1993 p96**). Evans and Hall (**1999 p7**) claim that we live in 'a culture which is pervaded at all levels by a host of cultural technologies designed to disseminate viewing and looking practices through primarily visually mediated forms'.

[9] Alan Turing proposed a phenomenological test for intelligence in computers, in which an observer corresponds with an unknown entity using a text terminal. If the entity is mechanical but the observer is unable to discern on the basis of its responses that it is not human, then it is is deemed to be intelligent (**Hodges 1983 p415-7**).

[10] I discuss later misunderstandings in the literature about the role of the retinal image.

Beginning with Panofsky (**1991** (**1925**))[11], some writers confuse this attitude to vision with one or more categories *within* visual realism. Lalvani is an example of the first confusion: 'That linear perspective or artificial perspective is a conventional symbolic ordering of space, operating within a cultural discourse of what constitutes the real, is supported by a number of observers. For instance, several anthropologists and social psychologists have promulgated the view that perspective is not innate but a culturally learned phenomenon' (**Lalvani 1996 p6**). Lalvani's 'for instance' elides two ideas: (1) that to choose to use linear perspective is to make a decision informed by cultural values and (2) that the perception of realism in linear perspectival images is similarly learned. This elision underpins the argument that visual realism cannot consist in any kind of pictorial correctness. The attack on such a kind of realism is also evident in the index of Crary's *Techniques of the Observer* (**1990**) which has no entry for *realism*, only for *'realism'* in quotation marks. Crary complains (***op cit*** **p32 n9**) at Wheelock's use of the term 'truly natural' in a discussion of the camera obscura, calling it a 'highly questionable notion', and also at Alpers' suggestion that it made possible a more truthful image.

By subsuming the straightforward question of pictorial illusion into the more widely accepted view that a *preference* for realist images is a cultural phenomenon, Lalvani is attacking the idea that for all observers, independent of culture, an image could fool them into mistaking a picture for a real scene. Two errors have contributed to mistaken answers to this question. The first is to give insufficient consideration to the conditions of viewing, and the second arises from the belief that realist pictures aim to match the *retinal* image.

The role of viewing conditions and context

Jones and Hagen (**1980**) ask 'What is it like for a person to see a picture for the first time?' as a way of discovering whether those accustomed to looking at pictures have learned a convention which allows them to interpret pictures correctly, or whether pictures are a natural mapping of vision. On balance they conclude (***op cit*** **p195**) that accounts of subjects having difficulty in 'attempting to make sense out of the shadings on the piece of paper' – a photograph – are based on flawed comparisons; they therefore tend to the view that the objects represented in pictures *can* be successfully perceived by image-naive subjects, suggesting that pictures are like what they depict in a simple sense which does not require cultural decoding. Oddly, however, Jones and Hagen do not bring into their argument the issue of viewing conditions, which is surely crucial to the success of the pictorial illusion. It would be hardly surprising that when a photograph is shown to subjects in a culture which does not use visual-realist pictures they might have difficulty in interpreting it. They are presented with a rectangular, delimited substrate whose tangible and visible reality as a piece of paper is clearly evident. It would be strange if the response to this circumstance were *not* culturally determined: image-using cultures expect to see segments of dislocated space carried around in wallets and displayed on hoardings and mantelpieces, but in an image-naive culture this decontextualisation will

[11] As Panofsky notes, relevant theoretical difficulties of depiction had been highlighted as early as the work of Leonardo, discussed later in this chapter.

naturally be a powerful counter-cue to any form of spatial perception within the picture. Failure to recognise an image in incongruous circumstances is not grounds for establishing the cultural conventionality of geometric perspective, photography, or any other picture-making technique. The context is decisive and as a result an important objective of any technology for creating pictorial illusion is to *suppress the observer's awareness of the context*. When I discuss film and related technologies in later chapters it will become clear that keeping the viewer's attention away from the artifice involved is a key device. One of the principal aids to the depth-spatial illusion in the Imax cinema for example is that the very large image extends beyond the boundaries of the normal visual field, depriving the observer of some obvious contextual counter-cues.

In 1425 Brunelleschi made a picture (now lost) of the Florence Baptistery which was allegedly remarkable for its match to the actual scene and which has been described as illustrating the discovery of the rules of linear perspective (**Edgerton 1976 p124-52**). While it is impossible now to test any claim that observers thought they actually saw the scene and not a painting, one can still ask whether it is theoretically possible that they might have done. It is necessary to imagine naive observers looking at various perspectival images, for example of the Baptistery, and to ask whether one spatial projection in the plane matches the scene in a way that all the others do not.

Several ways of controlling the context in which pictures are viewed have been tried historically in terms of the physical circumstances. Brunelleschi took three measures to support pictorial illusion. He closely controlled the geometry of the viewing conditions by forcing observers to look with one eye through a small hole, thereby disabling the perspectives of parallax (binocular perspective, motion perspective and the shift in amount of double imagery).[12] He also depicted a scene (the Baptistery) which was not only likely to be there in reality but which indeed *was* there to the prior knowledge of all his observers, so avoiding the incongruity problem noted above. Finally he avoided the falsity of static depiction of moving objects by using polished metal for the sky, reflecting real moving clouds.

The trompe l'oeil paintings of Gijsbrechts (Figure 4.12) also control viewing context, though by other methods. Subject matter is chosen which would normally occupy a shallow depth, such as a board with items pinned to it, or a collection of objects hanging on a door. By this means the painting's failure to offer convergence and motion parallax phenomena is, at least temporarily, unnoticed. Gijsbrechts always painted inanimate objects, avoiding the problem of stasis. His work also takes advantage of cultural conditioning in that the objects depicted are ones which viewers would not be surprised to find in such locations: for example, a boarded-up window in a wall. (To say that cultural conditioning helps to make the existence of the illusory objects more probable for particular observers is *not* however the same as saying that interpretation of perspective is itself culturally determined.)

[12] Parallax phenomena would in any case have been unimportant if all the parts of the scene were distant.

Figure 4.12.
Gijsbrechts: A Letter Rack with Christian V's Proclamation, 1671.
From Koester, Olaf 1999 *Painted Illusions: the art of Cornelius Gijsbrechts*, National Gallery, London, p23.

Figure 4.13.
Samuel van Hoogstraten 'Peepshow Box' late 1650s 58 x 88 x 63.5 cm; oil paint with some egg tempera and glue size; oak box.
The peephole is visible just to the left of the nearest vertical edge.
From **Cole 1992 p36.**

That Hoogstraten's boxes (Figure 4.13) exist at all is evidence of a fascination with visual realist ways of seeing which is part of a particular visual culture remarked by Alpers, but equally they are a demonstration of the possibility of PI-realism since they can confuse the viewer as to what is real space and what is illusionistic painting. Again this is only possible granted tight control of the viewing conditions, in this case by use of a peephole.

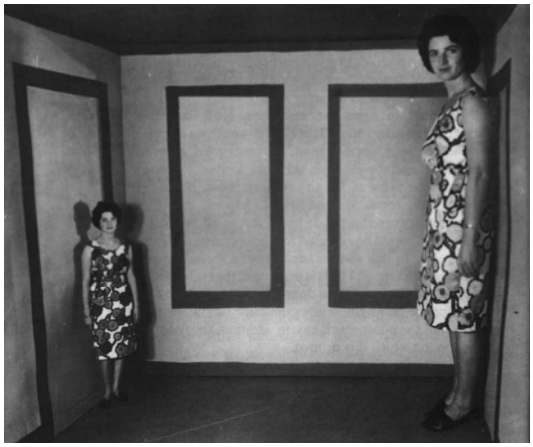


Figure 4.14. The Ames room.
Gregory's account of how this illusion works confuses cultural influences with constants of perception.
From **Gregory 1998 p186**

The Ames room illusion (Figure 4.14, described in **Gregory 1998 p185-7**) fools the observer because of culturally determined assumptions about the subject matter, namely that rooms are rectangular (an assumption which surprisingly is able to override the observer's assumption that people are all roughly the same size, and to

swamp any sensitivity to focussing phenomena). However, this culture-driven assumption about the room, a precondition for the illusion, is again not to be mistaken for proof of the culturally determined character of geometric perspective projection.

Under highly controlled conditions (especially when the absence of motion is suppressed), there is no *a priori* reason why an image, considered purely as a stimulus in the visual field, should not simulate the stimulus which would be provided (at least briefly) by the corresponding real objects. Any failure of the Brunelleschi, Gijsbrechts and similar illusionistic images to imitate the stimulus of a real scene does not reside in any weakness of their projective geometry, but in the failure of other perspectives, particularly binocular disparity and motion parallax, which cannot be dealt with by a static planar image in itself but only by extreme control of the conditions for viewing.

It might be objected that to control the viewing conditions in order to make the illusion work (even in theory) is 'cheating,' and certainly one could object to more extreme efforts intended to make picture and scene indistinguishable, such as diminishing the light levels, or making the subject matter itself planar, but, provided the necessity for these controls is admitted and not ignored, the argument for the possibility of PI-realism is still valid. There is no onus to produce a representation which could fool the observer in *all* circumstances since then it would have to withstand the effect, for example, of the observer looking at a canvas from the back, or taking off the VR headset and looking at it from a distance. The only objection can be to the *extent* of control which would invalidate the test of PI-realism, not the control as such.[13] Goodman for example considers Brunelleschian controls 'grossly abnormal'. However he seems to go too far in suggesting that such representations form an 'odd and futile argument for the fidelity of perspective' (**Goodman 1969 p13**): he is moving without sufficient justification from an attack on the conditions of the experiment, to suggesting that it is wholly invalid.

PI-realism does not imitate the retinal image

Some important figures have suggested that a picture represents the retinal image and this has produced considerable confusion. Gregory for example has said that **'***When an artist employs geometrical perspective he does not draw what he sees – he represents his retinal image* ... A photograph represents the retinal image – not how the scene appears.' (**1977 p174** *original emphasis*). His comment risks being misunderstood. What a PI-realist image resembles is nothing within the eye but the visual rays converging from the scene onto the eye.[14]

[13] This need for control of the test conditions is inevitable given that a representation is not the thing it represents. An objection to Turing's test is that it so narrows the opportunities for the observer to detect the illusion that it is an *unfair* test and therefore proves nothing (Searle 1977 p10-13).

[14] Gregory (personal communication, 17 May 2001) has clarified the intention of his remark as dealing with subjective perceptions of scale (discussed later in this chapter) in which objects are not perceived as having the relative sizes at which they occur on the retina because their apparent size is influenced by psychological considerations. He suggests that artists (when they are not using mechanical aids) tend to incorporate similar subjective scaling in their pictures. He is thus distinguishing the strictly optical from the psychological (as Klee did). However, it seems to me that referring to the optical pattern of the scene as the *retinal* image invites misunderstanding. His point about subjective scaling could more safely be expressed in terms of the external stimulus alone, without implying to the unwary that the image on the retina is itself accessible to vision.

Panofsky claims that 'artificial perspective' (that is, linear perspective in Gibson's taxonomy) is based on two assumptions: (a) that 'we see with a single and immobile eye' (an objection not in dispute here)[15] and (b) that 'the planar cross section of the visual pyramid can pass for an adequate reproduction of our optical image' (**Panofsky 1925** (**1991**) **p29**). The error of Panofsky's argument becomes clear when he says that 'perspectival construction ignores the crucial circumstance that this retinal image [...] is a projection not on a flat but a concave surface' (***op cit* p31**). So misled is he by this confusion between the stimulus and the receiving device that he suggests that we see lines as straight because of the conventions of painting and photography. His error is to consider at all the shape of the receptor which lies inside the eye. To simulate a scene, what is required is a perfect stimulus *outside* the eye, corresponding to the information in Gibson's 'sheaf of rays', to which the internal shape of the eye is irrelevant.[16] Alberti explicitly stated in 1435 that the operation of the eye itself is of no consequence to an understanding of his pictorial construction (**Alpers 1983 p53**). Though Panofsky was writing in 1925, the potential for confusion continues, as evidenced by Gregory's remark of 1977 repeated in later editions (**Gregory 1998 p184**).

Curved perspectives and 'Leonardo's paradox'

A second incorrect objection to linear perspective lies in another argument for curved perspectives. The ground for this argument is that if we face the buildings on one side of a street and look left the buildings get smaller with distance, that the same happens if we look to the right, and that the only way these impressions can be reconciled in the plane is by use of curved lines, as in Figure 4.15.

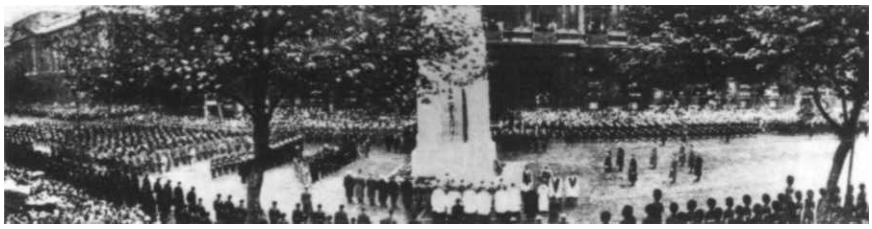


Figure 4.15. Panoramic photograph taken in Whitehall, London, from London *Times* 14 November 1966. From **Gombrich 1980** p197.

Such a supposition however ignores the role of the picture-plane as an interceptor of Gibson's sheaf of rays (or Panofsky's planar cross-section of the visual pyramid) which is *itself* at a variety of distances from the observer, depending on which part of the plane the user looks at. The picture plane is subject to exactly the same 'distortions' as the subject matter beyond (Figure 4.16).[17]

15 Richards (**1984 p4:8-4:11**) points out the need to distinguish between a fixed *location* for a normal eye free to move in its orbit and a truly *immobile* eye which exposes the retina to an unchanging stimulus, a distinction which, he notes, Goodman (**1969 p12**) fails to make.

16 The Turing test for artificial intelligence rightly does not take any account of *how* the mind of the observer works, but of whether that observer may be fooled by an external stimulus.

17 As indicated at the beginning of the chapter on Depth Perception and Depiction, it is generally assumed in this thesis that the picture surface is a plane orthogonal to the line of sight in the centre of the cone of vision. Almost all depictions are made on such surfaces. Nevertheless, as the example of panoramas and pictures applied to domes and other curved surfaces would suggest, there is no need for the putative surface which intercepts the 'sheaf of rays' to be planar. The Hoogstraten boxes offer a composite picture surface which is not even contiguous. In terms of *focus* rather than geometry, claims can be made for a spherical surface whose centre is at the eye, since all parts of its surface can be apprehended without the need for the eye to adjust to different distances. This is not an argument for using curved geometries in planar pictures.

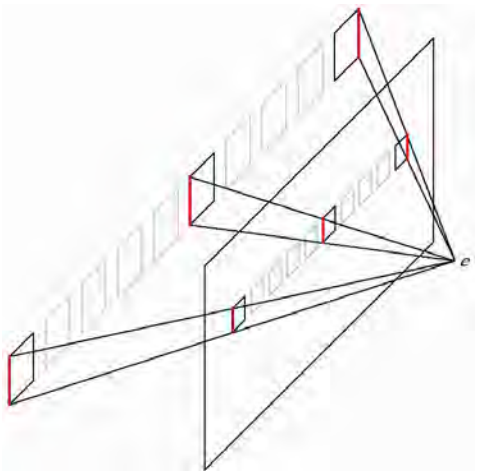


Figure 4.16.
The solution to the controversy over curved picture planes and to Leonardo's Paradox.
From the viewing station ***e*** the further a part of the scene is displaced from the central axis of vision, the further also from the eye is the corresponding mark on the picture plane. This corrects the alleged distortion. Other viewing stations for the picture will lead to incorrect proportions, but this is not an objection to the geometry of PI-realism.
Original drawing.

Dubery and Willats (**1983 p84-93**) make a related error when dealing with *Leonardo's Paradox*, an alleged problem of spatial representation first identified by Leonardo (Figure 4.17a-b). A row of identical and equally spaced columns runs across the observer's view and, using the rules of geometric perspective, the image of these columns is cast onto the picture plane. The result is, as Leonardo pointed out, a distorted image, since the columns at the ends of the colonnade (which are further away and therefore should be smaller) are manifestly larger in the image.

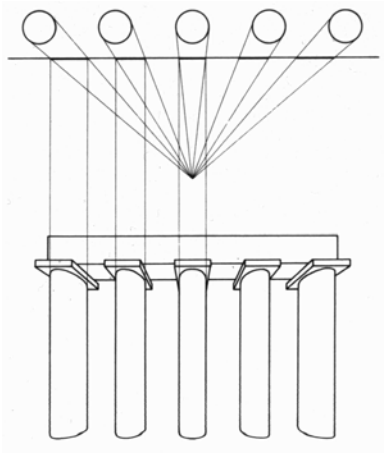


Figure 4.17a
A perspective projection of a row of columns, using 'normal artificial perspective'. The columns which are further away make a shape in the representation which is wider than the central column which is closest.

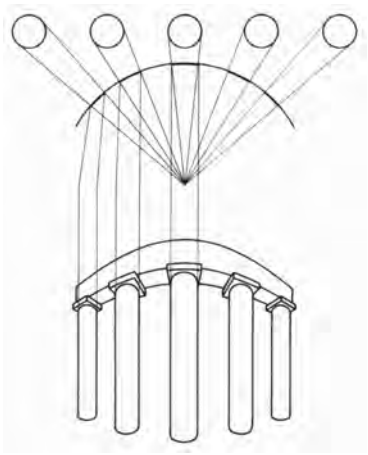


Figure 4.17b
The use of a curved projection to 'correct' the anomaly seen in R.16a

From Dubery and Willats 1983 p84 and 85

From this, Dubery and Willats deduce that the linear perspective system has no special claim to validity (***op cit* p86**). Unfortunately they also have forgotten that the image is itself viewed in perspective. That is, if the image is positioned correctly – at the same location in relation to the viewing station that a sheet of glass between the original viewpoint and the colonnade would have been – the further columns in the scene will also be further away from the observer's eye *in the image* so reducing their apparent size and restoring the correct proportions (Figure 4.16).

As Arnheim has pointed out (**1956 p233**), linear perspective *does* have a special relation to the visible world, since mechanically it can be a tracing of it. However, as

already indicated, the practicalities of enforcing the single correct viewing station are another matter. The misunderstanding of which Dubery and Willats work is an example arises because pictures are normally seen hung on walls or printed in the pages of books, where control of the observer's viewpoint is impractical. This has indeed created practical difficulties, evidenced in the work of Saenredam and Vermeer for example (**Dubery and Willats 1983 p86**), but this in no way negates the optical correctness of images when considered as a stimulus substituting, under controlled conditions, for a view of the real world. Strangely, Dubery and Willats acknowledge that a compulsory viewing position is the solution (***op cit* p87**), but rank this with choosing narrow fields of view and avoiding geometric subject matter as simply one contributory practical solution, rather than acknowledging that it undermines their case fundamentally.

As Alpers points out (**Alpers 1983 p244, n39**) Panofsky's view of the subjectivity of linear perspective has been attacked by Gombrich and Pirenne among others. Pirenne sums up the position:

> The picture in perspective of a scene or a set of objects is *not a replica of the retinal image* produced by the objects in the artist's eye. It is rather a substitute for the actual objects themselves, so constructed that it sends to the eye a distribution of light similar to that which would be sent by the actual objects, with the result that, for any given eye, the picture produces retinal images similar in shape and dimension to those which would be produced in the same eye by the actual objects.
>
> Pirenne, MH 1948 Vision and the Eye, London, p15
> quoted in **Edgerton 1976 p163**, *emphasis added*

What is required is to organise the pictorial surface to stimulate the retinae in a way which matches looking at a scene, and control the conditions of viewing to ensure that the illusion is not undermined.

At this point it may seem that this special correctness of linear perspective counts fatally against the main arguments of this thesis in which I frequently suggest that the construction of visual representations is a pragmatic matter of making rather than matching: now I seem to be endorsing the opposite position. However the ongoing argument hinges not solely on the claims of geometrical correctness of PI-realism, but on the impossibility of deciding other aspects of the truth to be depicted (introduced in the previous chapter) and the practicality and usefulness of PI-realism. These last are now considered.

6 Difficulties of PI-realism

For convenience I shall from this point call PI-realistic those images which are designed with a view to *approaching* true pictorial illusion, even though as discussed they can only offer an illusion of natural vision under highly controlled conditions.

Difficulties of attaining PI-realism

One historical explanation of why not all pictures are PI-realist is that image-making cultures other than our own (that is, from the Renaissance onwards) were incompetent either in terms of perception or execution. In Hogben's commentary on oriental art noted earlier, there is an assumption that societies prior to, or geographically

remote from, the Renaissance did not make PI-realist images because they were unable. Renaissance literature tends to support this view by particularly emphasising the issue of the competence of various artists in constructing illusory spaces and objects. Thwaites belongs in this tradition, suggesting that 'Greek and Roman art reveals that artists had a *vague notion of perspective...*' (**Thwaites 1999 p222** *emphasis added*). Similarly he suggests that 'the creators of the paintings at Lascaux and Altamira used a *primitive method* of portraying distance: they drew figures in different sizes...' (***op cit* p221** *emphasis added*).

In opposition to this argument, cultural theorists such as Crary and Lalvani have argued that competence is not the key issue. This is where the concept of visual culture is useful: it highlights the possibility that the construction of PI-realist spatial illusions may in many cultures have elicited little or no interest. Given the widespread sophistication of both artistic and technological practice in many cultures, and the fact that their pictures often come close to (but do not become) PI-realist images, the balance of probabilities is surely in favour of this view that cultures had (and have) other concerns, turning their attention away from this supposed 'goal', rather than that they were unable to reach it. After all, though our own society is noted for its ability to produce PI-realist images, it also produces and uses innumerable kinds of pictures which are not of this kind, without the makers being accused of incompetence. This alone suggests that there are other reasons to make images which fall short of, or substantially diverge from, PI-realism.

This argument could lead to a view that picture-makers in all cultures *could* have made PI-realist images but chose not to. However this view is not easy to defend. It implies that making such images is a fairly straightforward and natural activity and therefore something likely to occur in any image-making culture. This seems unconvincing: although I have been arguing that the PI-realist picture has in one sense a uniquely 'natural' relation to the viewing of actual scenes, I do not argue that the *making* of such images is equally 'natural'. On the contrary, the making of PI-realist images seems distinctly difficult. A simple characterisation of the problem, that it is easier to draw what one knows than what one sees, seems borne out by the problems experienced by novices in drawing.[18] The errors which occur (for example, incorrect estimation of the trapezoidal shape of rectangles when seen obliquely) are just as one would expect when the picture-maker is having great difficulty in suppressing the knowledge of how things 'really are' – that is, how they look when seen from other angles such as canonical points of view[19] and how they seem or would seem to the sense of touch. Though the objective PI-realist image convinces when we see it, it is something impossible for the picture-maker to apprehend pre-pictorially by introspection (to look at the retina from within, as it were). This seems to explain the proliferation of techniques and devices (the camera obscura and Dürer's machines for example) which aim by external means to aid the capture of the PI-realist image.

[18] These observations are partly based on my own experience in teaching drawing over many years.

[19] There is evidence that some views of objects are more significant than others, probably because they differentiate the object with greater efficiency, summarised in **Bruce, Green and Georgeson 1996 p224**. Canonical views are further discussed below.

Practical difficulties of using pictures to simulate vision

I have already noted how any shortfalls of picture-making from full optical realism must be concealed in order to achieve a true illusion of looking at a scene. If a binocular system is used, then binocular perspective and shifts in the amount of double imagery at contours are achievable.[20] However, time-based cues (those dependent on motion or on the dynamics of the visual system) are difficult or impossible to solve in a static picture (I have noted how Gibson came to consider the absence of motion in pictures to negate any attempt to replicate vision pictorially). What cannot be achieved in the image itself may be compensated for by controlling the content and the context, the conditions under which the image is seen, but there will usually be real practical difficulties in enforcing these conditions, in which case the realism is no longer one where the image can be mistaken for an actual scene. Every other situation is 'second best' in which, whatever form of visual realism is in use, it is *not* defined by illusion as such.

Other practical difficulties arise from the nature of the media in use. These may involve a minor shortfall from optical verisimilitude, for example the smaller tonal gamut of a picture compared with natural vision (though this may be compensated for by the comparative rather than absolute nature of many aspects of the visual system). Also there may be a fundamental lack of capability in the chosen medium, as for example in pencil drawings or wireframe computer graphics. This problem is often side-stepped by the use of what I have named *illicit marks* – marks which have no corresponding presence in the scene. The 'halo' of yellow around Dürer's figure is an example (illustrated earlier in Figure 4.06), but perhaps the commonest is the drawn outline. So ubiquitous is it in pictures of all kinds that one could easily forget that it has no basis in the scene. What does it depict? It seems to stand for an aspect of the world *after* it has been perceived, after the scene has been processed by the visual system: it is certainly incompatible with PI-realism, which must be confined to a wholly external stimulus to vision.[21] The use of outline is one of many pictorial attempts at VE-realism, the presentation of a stimulus which in some degree evokes the *visual experience* of looking at a scene, not simply the optical pattern of the scene which is looked at. I discuss such illicit marks further below.

Practical difficulties are not in themselves sufficient to account for the widespread creation of pictures which diverge from PI-realism. It is necessary now to consider ways in which PI-realism may fail to fulfil the functions intended for a depiction.

Functional mismatch of PI-realism to the objectives

There are many reasons why the closest possible approach to PI-realism may simply not be suited to the task in hand. Even since the advent of photography, there has

[20] However, Vince (**1995 p53 and 59**) points out that while in natural vision the eyes converge on a point of interest in the scene and the location of the doubled imagery is therefore affected, in binocular computer 3D simulations the two views look out parallel to each other: 'Ideally, the geometric model used for computing the stereo pair requires to know the fixating point, then it can rotate the virtual eyes and their respective projection planes to mimic convergence.' This may be why the depth seen in binocular pictures often seems false, as though it were a succession of planes rather than fully volumetric.

[21] Marr's model of vision (**1982 p37**) proposes that mental constructs equivalent to outlines are a primary means of segmenting the scene at a very basic level, but this does not alter the fact that the lines are post-optical, not in the scene.

been a continuous outpouring of images which do not use the spatiality of PI-realism. The representations favoured by architecture, product design, engineering, magazine advertising and many other image-using trades and professions use convergent perspective only as one among a range of possible projections. This is partly because it is not always as expressive in informational terms as the alternatives. In place of linear perspective, orthographic, axonometric, isometric and hybrid projections offer benefits such as constant scale or greater clarity about the interrelationship of multiple viewpoints. In the previous chapter, many disparate reasons emerged why the full depth cues of natural vision might be suppressed, distorted or subverted in order to convey information more effectively.[22]

In addition to these objections on the grounds of informational function, there are reasons to question PI-realism when the image is considered as part of a process. As Baker (**1993 p30-1**) emphasises in relation to computer graphics, pictures perform a role in personal and social activity. Sketches for example are used for externalisation and reflection, as well as communication to others; technical drawings to convey unambiguous instructions to others; highly pictorial renderings to persuade colleagues, clients or planners. An image which closely matches natural vision has by its nature a character of completeness and non-negotiability (**Schofield 1996**), and several authors have pointed out the benefits of avoiding such characteristics in some circumstances. When a sketch is an element in a process, it is clear that there are merits to imprecision and tentativeness in depiction. Arnheim (**1993**) calls sketches 'guiding images' whose role as externalisation facilitates the design process. As long as the guiding image is still developing, the sketch remains tentative, generic, and vague and this vagueness is not a negative quality. The sketch stands for a whole range of possibilities without being wholly committed to any one of them. This is a striking example of the concept of affordance in which the representation is chosen because it is *less* dictatorial of the mental visualisations which the user derives from it than in a more fully worked out picture. Scrivener and Clark (**1994 p98-9**) also suggest that the sketch operates by affording multiple mental images rather than being confined to one, and that the lack of specificity evokes greater imaginative work on the part of the perceiver. This is an argument for the limitation of visual realism in the interests of expressiveness. Their characterisation, like Arnheim's, is essentially an operational one, in which the picture-maker puts down marks which afford interpretations, a process quite different from mapping Gibson's 'sheaf of rays' or Pirenne's 'distribution of light'.

Another departure from simple realism is that of the illustrator who suppresses parts of a machine in order to reveal concealed mechanisms (Figure 4.18). The whole is a hybrid of two views, one a lively portrait of a car in action, the other a static depiction of its anatomy. The approach to realism is driven by the purpose.

For the cutaway car, the inclusion for informational reasons of normally invisible segments of the subject matter creates considerable extra work for the picture-maker,

[22] The word 'reasons' might seem to imply that these decisions were always deliberate – perhaps even a personal decision by a particular picture-maker – but this is not intended, as explained elsewhere.

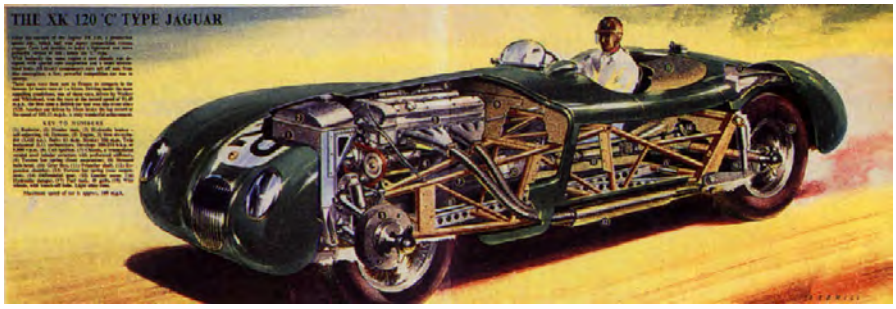


Figure 4.18. Cutaway visualisation of a car, 1952.
From Sabin, Roger 1996 *Comics, Comix and Graphic Novels: a history of comic art* Phaidon, London p47

but in other cases informational demands and implementational constraints may coincide. In Figure 4.19 implementational constraints dictate that size perspective is suppressed since this enables views to be pre-imaged rather than calculated on the fly from model data. However such a projection also has the advantage of ensuring that all parts of the city are equally resolved and that the foreground is not unnecessarily favoured. As computer power increases, implementational constraints will decrease, but the functional argument will not. Nevertheless, a search for a greater sense of realism may on balance cause a change of practice: users may *prefer* the realism of convergent geometric perspective. It is important to note that such realism would not be informational, but affective: it is concerned not with giving the user more information about the scene, but a different relationship to it.

Figure 4.19.
Screen-capture of **Sim City 2000.**
© 1993 Sim-Business.
Published by Maxis, London, 1993

Mismatch of PI-realism with symbolic functions

An interesting issue of realism is raised by Richards (**2000**) who suggests a three-dimensional relation between types of graphic communication (though here I deal only with two of his dimensions[23]). These are the *mode of correspondence* ranging from the literal to the metaphorical, and the *mode of depiction* ranging from the abstract to the figurative (Figure 4.20)). Richards describes these as 'independent variables' (**Richards 2000 p97**) which is clearly correct in the sense that the picture-maker may alter each without altering the other, but I suggest that to the viewer or user of a graphic the mode of depiction *affects* the perceived mode of correspondence and that

[23] The other dimension is concerned with *modes of organisation* such as grouping and linking.

this represents another of the limitations on the functional usefulness of strong forms of realism. Pedersen (**1998**) found that imagery with a high level of realism caused symbols to be taken literally. Strong figurativeness tends to denote strong correspondence: thus if users see a picture which looks very like, say, an engine, they are more likely to interpret the configuration of the parts as showing how the parts really are in an actual engine than if there is low figurativeness leading them to think that it is a diagram or a symbol and should not be taken literally. If the London Underground map looked like an aerial view of a city – was towards the realist pole of the axis of depiction – users would assume that the map was more topographical than it actually is (they would be estimating its position as near the literal pole of Richards' axis of correspondence). The role of imputed intention complicates the issue: a photograph and a highly realist non-photographic picture will be interpreted differently. The photograph will be regarded as *accidentally* specific (and irrelevant specifics therefore be ignored), whereas the other will be regarded as *intentionally* specific and its detail will be taken to be significant. As always, content, context and depiction interact.

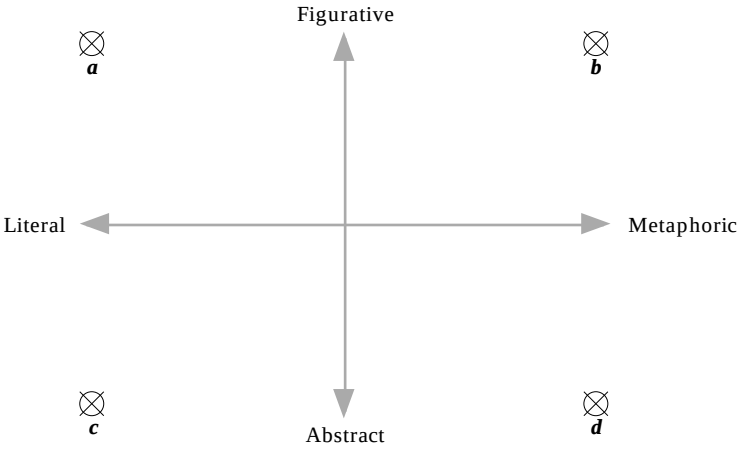


Figure 4.20.
Modes of correspondence (horizontal) and modes of depiction (vertical) in graphic representation.
Derived from **Richards 2000**, p97 (The original model is three-dimensional, the axis omitted here being 'modes of organ-isation'. The label 'metaphoric' is 'non-literal' in the original, while 'non-figurative' in the original is here called 'abstract').

Using the example of depicted trees, the extremes represented by ***a***, ***b***, ***c*** and ***d*** can be considered as follows:

a an image with strong visual resemblance to a tree and intended to convey information *about* trees, perhaps an illustration in a field guide. There is no doubt that it is a *picture*, within the definition used in this thesis.

b a perhaps identical image of a tree, but one which is intended as a metaphor for, say, the major divisions and sects of a religion. It is not a picture of the religion but is still undoubtedly a picture of a tree.

c an image which is intended to convey information about trees, but not through strong visual resemblance to an actual tree. For example it might diagram the flow of nutrients through the parts of trees. It is likely that in the nutrient-flow diagram the top of the tree is at the top of the diagram, but bark, leaves, etc. will probably be absent. This is perhaps still within the definition of a picture.

d an image which might exactly resemble the diagram at ***c*** but which is about, say, the flow of messages through a computer network. It is not a picture (unless the branches of the 'tree' are organised to correspond to the actual spatial relations of the network's parts, in which case it ceases to be in any useful sense a picture of a tree but starts to become a picture of the network).

Hogben (**1949**) suggests that the principal gains from the use of geometric perspective accrue to scientific communication, where the ability to convey a precise and relatively unambiguous depiction of what things look like to an original observer (for example in a study of a dissected body) is of extreme importance for the development of empirical science.[24] For him this is of greater functional benefit to science than it is

[24] It is ironic that even realist pictorial images are (theoretically) infinitely ambiguous yet they have in practice been important for their informational expressiveness. This arises in part from the fact that the things depicted are generally not abnormal objects (this greatly decreases the range of possibilities of what may have been depicted) and from the fact that optimal views are chosen to maximise informational expressiveness. This theme is taken up in the next chapter in relation to film.

for symbolic activity, for which he acknowledges the third dimension is an aspect of the perceived world which may conflict with other needs and purposes. Ivins (**1953 p127**) notes the significance of the accurately repeatable illustration: '...without prints we should have very few of our modern sciences, technologies, archaeologies, or ethnologies – for all of these are dependent, first or last, upon information conveyed by exactly repeatable visual or pictorial statements.' Edgerton (**1976 p164**) similarly argues that without the conjunction of printing and linear perspective the whole subsequent development of modern science and technology would have been unthinkable. Arnheim, following Ivins, suggests (**1956 p233**) that the contemporaneous emergence of prints as reproductions of a graphic original and pictures as reproductions of the world is not accidental. He regards the importation of a 'scientific criterion of correctness' into the visual arts as 'a dangerous moment in the history of Western thought' (***ibid***).

Marr and Nishihara (**1978**) describe vision as 'a process that produces from images of the external world a description that is useful to the viewer and not cluttered by irrelevant information.' This strictly functionalist characterisation of vision transfers poorly to the uses of pictures, because it leaves no role for the affective uses of pictorial realism. It could account for the realism of a scientific illustration, but not for that of a holiday photograph, since the latter seems to be full of 'irrelevant information.' Birrell, arguing (**1999 p328-9**) against the obsession with visual realism in computer graphics, suggests that 'we should consider virtual objects by what they are and what they do, not what they look like' but this is to beg the question of what depictions are for and to reject without reason the mimetic representation of visual scenes. The question remains, What is strong realism for, when it is *not* intended to fulfil some practical purpose?

Baker has described the aim of realist computer graphics as being to map a three-dimensional world to a plane surface in a '*convincing* and *effective* manner' (**Baker 1993 p33** *emphasis added*) and this seems exactly right. In most cases either the demand is an operational one of effectiveness, in which case the realisms and concomitant spatialities chosen are those which support the practical functions which the pictures serve (as for example in the military training applications cited earlier), or the pictures must merely convince. In the former case complete truth is not required (even if it were achievable) since only those realisms required for the task are called into play, and in the latter case truth is not required at all – it is not necessary that something be true for it to be convincing. The requirement is an affective one, not informational. Such images must evoke in an *apparently* natural way the subjective character of lived experience. This will turn out also to be the criterion for most film-making.[25]

[25] In dealing with these affective issues, the position is complicated by the user's ambivalence about illusion. An awareness of the fact of illusion is in many cases itself a part of the user's pleasure (as in Coleridge's 'willing suspension of disbelief', **Coleridge 1817 ch.14**) an appeal like that of magic in that it creates something from nothing and is able to transmit pieces of reality from place to place.

It is fundamental to the concept of 'convincingness' that realism does not reside solely within the artefact. The viewer brings to a picture an awareness of the context in which it is displayed and the technology which was used to make it, and this also leads to many different kinds of 'being convinced'. When Sontag says (**1977 p86**) that 'the flat, usually rectangular, images which are photographs make a claim to be true that painting can never make' she is referring to conviction about what is depicted in which the awareness that the medium *is* photography alters the context irrevocably. While strict PI-realism can be defined purely in terms of the relation between the image and the scene, all other forms of realism must take into account the context, including the viewer.

For convincingness, PI-realism is not necessary. However, if a sense of 'just seeing' the scene is to be promoted, any conscious awareness of encoding or convention needs to be suppressed. Crucially, this does not mean that convention is not there, but that it is not noticed. In this sense the simulation of unmediated viewing may be better defined by absence than by presence. It is the *lack* of a sensation of encoding or convention in the image which makes a picture seem close to vision. This leads back to consideration of how the spatiality of pictures may be tailored to offer a more broadly defined *visual experience*.

7 PI-realism compared to Visual Experience

I have enumerated numerous difficulties which stand in the way of PI-realism, though acknowledging that technologies such as those of virtual environments have the potential to eliminate many of them, and have drawn attention to the functional purposes which pictures may serve, for which strong visual realism will only in some cases be appropriate.

It is necessary now to contrast the PI-realist picture with the alternative approach to realism which I named earlier: VE-realism, which aims to evoke *visual experience*. Since the latter aims to capture subjective aspects of the situated experience of seeing, and since those aspects will be chosen differently in the light of whatever the picture-maker intends, there clearly can be no one VE-realist picture which corresponds truthfully to the scene. VE-realism can therefore only be defined in terms of a general approach, while PI-realism by its nature could be defined once-for-all and explicitly.

I will discuss three drawbacks of PI-realism considered as a possible means of conveying visual experience: (1) its projection system (even though I have argued its optical correctness); (2) its relationship to time; (3) its being limited to the imitation of retinal stimuli. I deal with each of these in turn. A point which will emerge concerns the relation which I introduced earlier between the model or scene ***M***, the view ***V*** and picturing ***P***. It will become clear that only in the case of PI-realism can the clear-cut distinction between them be maintained.

PI-realism's projection system

There are grounds for selecting alternative projections when the intention is to

capture the *subjective* experience of a space. I have already indicated how the 'flattened' depth of oriental landscape painting promotes a different kind of exploration of the landscape. The form of such paintings might in some ways replicate more accurately the visual experience of looking at landscapes. For example, the phenomenon of subjective size constancy has been noted by Gregory (**1977 p174**): 'distant objects look too small in a photograph – it is a common and sad experience that a grand mountain range comes out like a pitiful row of molehills.' When looking at an object in a scene there is a tendency to scale it to a size nearer to its local size, so that features on the horizon are effectively brought nearer. There seems good reason to suggest that the particular spatial construction of the oriental painting might – in this respect – have a better match with the *experience of space* than does the PI-realist image.[26]

Figure 4.21.
Van Gogh: Vincent's Bedroom at Arles, September 1889
From Uhde, Wilhelm 1972 *Van Gogh*, Encyclopaedia Britannica, London, Plate 48.

Dubery and Willats (**1983 p92**) claim that Van Gogh may have painted his room in the way he did (Figure 4.21) as the result of the use of a measuring-stick such as a paintbrush or pencil in capturing a wide-angle view, which will always tend to result in a curved perspective. Their view assumes that this distortion took place as it were under Van Gogh's nose without his noticing it.[27] However the most convincing explanation for the painting's appearance is that the painter was not trying to capture the optical image of the room but to make an artefact which *affords some particular visual experience*. Viewers 'sense' their own location in relation to the scene: the room feels like a small, intimate space because of the 'distortions' imposed on it. Whether the artist reached this result through the initial use of a measuring device producing a result which he liked and developed, or whether the whole image is constructed from the outset with the effect in mind is, for current purposes, irrelevant. The point is that the final decision was a pragmatic one based on an assessment of the effect on the observer.

[26] A similar dilemma arises in relation to *colour constancy*. When an object in a scene is observed – say a white tablecloth under trees on a sunny day – several factors causes it to be seen as white. But in the scene the tablecloth may actually appear quite dark and have a distinct non-white hue, a phenomenon that particularly captured the interest of the Impressionists. Which colour should the tablecloth appear in a picture? The 'right' answer will depend on the purposes of the picture-maker and the expectations of the picture-viewer, partly formed by the experience of other pictures.

[27] Dubery and Willats' theory is made more questionable by their admission that Van Gogh made other pictures of his room using both more and less curved projections (**Dubery and Willats 1983 p122 n8.5**).

The relationship of space to time

Under this head are subsumed many different ways in which the experience of a PI-realist picture may fail to accord with the experience of space considered over time. Willats offers a summary of the problems of relating what I have called PI-realist images to subjective experience:

> But in fact photographs, or pictures which look realistic in a photographic way, only give one kind of truth: truth to appearances. Another kind of truth, equally important both to artists and to architects and engineers, is truth about the shapes of objects as they really are, independent of any particular viewpoint. This is the truth that Cézanne and later the Cubists were after: the kind of description of objects that we arrive at in our minds after the visual system has processed and collated the immediate and transitory sensations available at the retina. David Marr, attempting to describe this end-point or goal of the visual process, called it the '3-D model' and described images of this kind as 'canonical'.
>
> **Willats 1990 p237**

This summary conflates three things:

- 'truth about the shapes of objects as they really are, independent of any particular viewpoint'; this is the problem of representing what is *known* about an object rather than what can be seen at a single moment from a single point of view: for example indicating aspects of the further side of a body which are not actually visible from a particular view.
- canonical views: these have some overlap with the 'truth about the shape of objects' of the previous category but they are a subset of it in that they represent the most discriminating or iconic view: for example drawing a table-top more like a square than a trapezoid.
- 'the truth that Cézanne and later the Cubists were after' which is a synthesis through accumulation of multiple particular, discrete observations.

In the previous chapter I discussed another aspect of time's relation to pictures, the representation of cues such as changes in focus arising from the time-based operation of vision. There is yet another aspect which I acknowledged could not be directly captured in a static picture, namely motion in the scene, which I deal with now.

Capturing some aspect of motion in the scene

For the pioneers of photography, instantaneity was a goal[28] – but vision takes place as part of continuous experience. Early commentators on photography were intrigued by the way the long exposure times of the period seemed to empty the streets of moving figures: anything which travelled through the space became invisible (**Frizot 1998 p28**). We 'know' that a person who moves through a space is a coherent form and not a ghostly impression stretched through the environment. But when something moves, how should it be represented? Between the waterfall which might in an early photograph be a white streak or later a frozen crystal sculpture, which is closer to visual experience? This seems a problem which the moving image of film can solve, but to which the static picture cannot offer a definitive answer. Any photograph has a particular relation to continuous sense-impressions and so is an

[28] Nièpce's thirty minutes in 1829 gave way to Nadar's approximately 20 seconds in 1860, so that Disdéri could write 'What remains to be done, I think... is to speed up the process further: the ideal solution would be to obtain instantaneity' (**Virilio 1994(1988) p21**) and by 1878, snapshots of 1/25 of a second were possible (**Gautrand 1998a p233**).

artefact contrived to afford certain aspects of visual experience. It is designed to evoke *selected aspects* of natural vision, both less and more than could be apprehended by vision itself in the same brief time.

Other forms of constructed image employ different analogies to the workings of the human visual system. One of the crudest attempts to make space stand for time appears in the work of the Futurists and in some works by Duchamp. Probably influenced by multiple-exposure photography, they painted multiple 'echoing' images in a single picture to show the successive locations of a moving object. The most interesting aspect of these is that they seem quite unacceptable – unrealistic – as 'natural' representations of visual experience: the element of 'encoding' is highly apparent. In this they contrast strongly with both the frozen instant and the smeared impression of the photographic styles just discussed, either of which seems much nearer to naturalism. Part of the reason perhaps lies in the collapse of automorphism in such images. Object constancy in natural vision ensures that a single moving object is apprehended as just that, a single object. In these 'multiple-exposure' images however the object is multiplied. This draws attention to the artifice of picture-making which breaks any illusion of 'just seeing'.

A related spatial device is the use of two or more separate images of the same person or thing in a representation. Figure 4.22. shows a part of a page (drawn in 1340) in which Noah appears twice.

Figure 4.22.
The Velsilav Bible: Noah Building the Ark, in a Bible made in Prague circa 1340.
From Unger, Richard W. 1991 *The Art of Mediæval Technology: images of Noah the shipbuilder*, Rutgers University Press, New Brunswick NJ, plate 31.

Cole (**1992 p9**) notes that Figure 4.23, which shows a protagonist at two different stages in a narrative, is on the cusp of two different approaches to representation: the mediaeval convention of combining two episodes of a story in one scene and the early Renaissance fascination with new perspective techniques. The first can be thought of as representing informational expressivity – economically conveying several times in one image – while the other gives priority to the relation between the picture and the viewer – the affective aspect of picture-making. While the later history of painting shows the dominance of the latter, pictures with other purposes continue to use the informationally more expressive form: illustrations of assembly of machinery for example routinely show multiple time-states (Figure 4.24).

Figure 4.23.
Giovanni di Paolo: St John the Baptist Retiring to the Desert 31x39 cm; egg tempera on poplar.
From **Cole 1992 p9**.

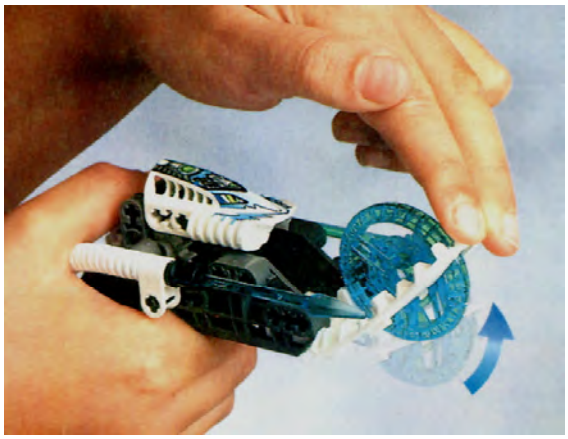


Figure 4.24.
Illustration of the operation of a toy.
Two time states in one image.
From instruction leaflet for Lego Technic product no. 8511, © 2000 Lego Group, p19.

The issue is one of *apparent naturalism* in which elements unacceptable in the kinds of picture which imitate natural vision *are* accepted in pictures where overt 'encoding' is considered reasonable. What matters is not that realism should be complete but that it should lack any *evidence* of convention.

Representing what is known about an object (rather than a single moment and point of view)
Alberti asserted that 'the painter has nothing to do with things that are not visible. The painter is concerned solely with what can be seen' (Alberti 1435, quoted in **Park 1997 p135**).[29] Hogarth acknowledges the difficulties of reconciling what is seen and what is known, in his practical advice to new painters:

> in the common way of taking the view of an opake object, that part of its surface, which fronts the eye, is apt to occupy the mind alone, and the opposite, nay even every other part of it whatever, is left unthought of at the time: and the least motion we make to reconnoitre any other side of the object, confounds our first idea, for want of a connexion of the two ideas, which the complete knowledge of the whole would naturally give us.
>
> Hogarth 1753 *Analysis of Beauty* London, First Edition, p8 from **Podro 1998 p111**

In the Albertian tradition Gombrich (**1980 p197-200**) proposes a test for the validity of linear perspective based on its *omission* of anything which cannot be seen from a single viewing station.

[29] Park comments on the radicalism of Alberti's statement in contradicting the Scholastic tradition in which 'visible form is only one index, often faulty, of the true nature and substance of whatever one wishes to portray' (**Park 1997 p135**). This is just the kind of change in cultural attitudes to vision which the concept of visual culture usefully captures.

It is clear that the ability to give more form, structure, depth and so forth than PI-realism allows, in short to be more expressive, encourages many kinds of picture-making which are not limited to what can be seen in a moment from a single viewpoint. This also helps explain the multiplicity of different kinds of images, the 'sensory chaos' of the world of pictures (**Hagen 1980 p9**), since each type of picture prioritises the affordances which are best suited to the particular aspects of form which it is intended to convey. Picture-makers often study their subject from many more angles than they depict, or investigate its underlying structure. This is a tradition which connects painters such as Leonardo and Stubbs to computer graphics researchers such as Waters (**1987**) or Scheepers *et al* (**1997**). It is not necessarily incompatible with a PI-realist image: the underlying model may simply lead to a more accurate mimesis of appearance. However, this merges into a kind of illicit mark-making where aspects of this other evidence are actually incorporated in the picture.

Faced with the limitations of PI-realism, picture-makers have invented numerous, variant techniques for incorporating in pictures what is known as well as what is seen. Yet at the same time they have generally retained the basic automorphism of visual realism. VE-realist picture-making pretends to be 'like seeing,' but the seeing which it resembles is augmented by the pragmatic incorporation of data derived from multiple views.

Capturing canonical aspects of objects in the scene

This is a form of VE-realism which often tends toward the symbolic or iconic – though still undoubtedly pictorial – and some examples might not be considered 'realistic' at all. What matters here is 'simple object identity without attention to momentary appearances' (**Hagen 1980 p13**) – not so much the capturing of time-based perceptions in pictures, but the elimination of time. I have already noted that in strongly realist styles it is possible for a representation both to picture something particular and stand for a generic concept, but it is a fundamental aim of many visual representations that they should eliminate the particular. Such pictures frequently adopt canonical views and suppress accidentals, for example in the images used for road signs, airport signage and icons in the digital interface. To achieve such images, omission from the optical image is not in itself enough: the scene or object must itself be configured in a canonical state. Not only ***P*** and ***V***, but ***M***, must be designed for the purpose.

Ethnographers have observed how in the case of magic carried out on images, the images do not display any likeness to the man who is bewitched: there is nothing resembling a portrait, merely 'a very schematic representation, a poorly executed ideogram' (**Taussig 1993**). These images are recognisable as people, because a canonical pose is depicted. Such images have *just enough realism* (very little) to afford the operation for which they are designed. The maker of the magic identifies the representation with a particular individual and this individual therefore does not

need to be visible in the representation itself. Model, view and context (including the user) taken together comprise a sufficiently expressive system.

The idea of the 'most characteristic aspect' helps to explain representations such as children's drawings which seem to negate the need for a concept of a viewing station altogether – as though the transformation was ***M*** → ***P*** rather than ***M*** → ***V*** → ***P***. If there is automorphism it is not of the kind discussed so far: now it is the connectedness of objects which is represented visually and not their spatial relation when seen from any place. The topology of what is depicted begins to dominate its topography. Burton has developed a computer simulation to explore some of the processes he believes are involved in children's drawing (Figure 4.25). He contrasts his algorithms (in the program *Rose*) with the process which would define a viewpoint on an object: 'A projective interpretation would describe the human as drawn as seen from the front, and the horse as seen from the side. It is important to realise that for Rose this is not the case. Rose uses the secondary axis to identify the *most informative way* to differentiate between opposite sides of an object' (**Burton 1995 pC167**).

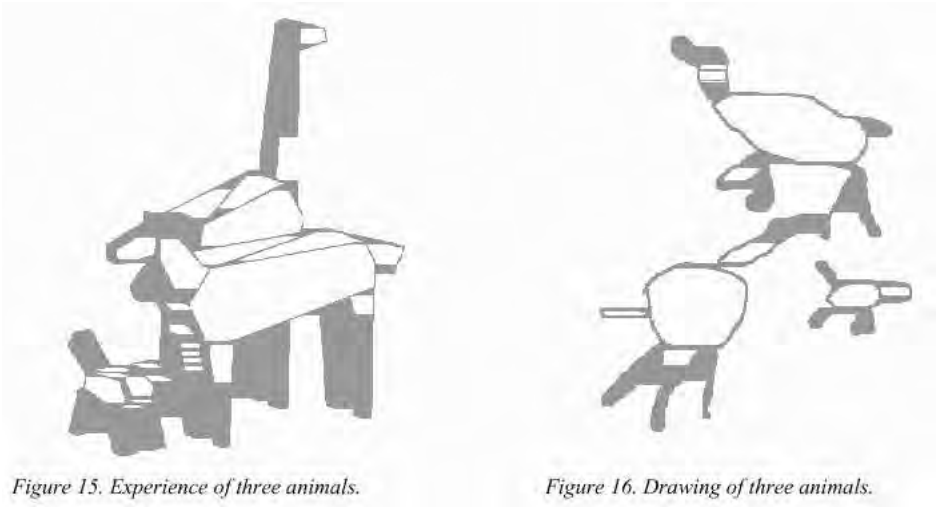
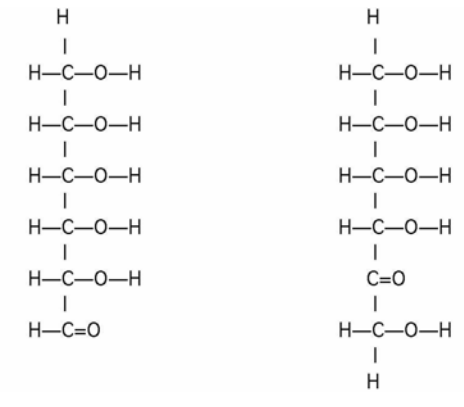


Figure 15. Experience of three animals. *Figure 16. Drawing of three animals.*

Figure 4.25. A viewpoint-based picture and a viewpointless view of three animals. Burton rather misleadingly captions the first 'experience of three animals' a title which would probably be better applied to the second drawing. From **Burton 1995** pC169 Figs. 15 and 16.

Drake (**1986**) discusses the balance of picturing and non-pictorial methods in conveying scientific information. In Figure 4.26, as with many diagrams having visual



Glucose Fructose

Figure 4.26. Diagrams of glucose and fructose molecules. Redrawn from **Drake 1986 p153**.

referents, a pragmatic combination of the pictorial and the schematic is used. A disadvantage of a more pictorial representation such as a 'ball and stick' model rendered as a picture would be that it invites the user to misinterpret the image as though the 'ball and stick' were not metaphorical but literal (as in the interference between Richards' modes noted above). In this more schematised representation it is highly unlikely that the user will make this error. Yet it is still able to benefit from some of the 'just seeing' characteristics of pictures. As Drake puts it, these diagrams make it easy to 'find one's way about in a molecule, so to speak' (**Drake 1986 p153**).

Representing multiple aspects of a space derived from a series of observations

While the previous section touched on an enormous population of images, the representation of multiple particular aspects in a single image accounts for far fewer examples.

In painting, the aggregation of multiple views is important in the work of both Cézanne and the Cubists, yet there is little resemblance between them. The two interpretations of the same idea are perceived in very different ways, one being seen as pictorial, the other as distinctly schematic.

Figure 4.27.
Paul Cézanne: Vase of Tulips, Oil on Canvas, 1890-94.
From Elgar, Frank 1969 *Cézanne* Thames and Hudson, London.

Cézanne's *Vase of Tulips* (Figure 4.27) has a superficial resemblance to a conventional linear perspective picture. Yet Cézanne rejects the depiction of a single moment implied by PI-realism. For example, the two segments of the far edge of the table fail to 'line up'. The edge of the table at left is defined in relation to the left edge of the vase, that on the other side in relation to the right.[30] Cézanne frequently explores such anomalies because he aims to use the planar image to recreate the exploratory visual experience of looking at a three-dimensional space. The spatiality of Cézanne's work is designed to allow him to articulate the aspects of visual experience which interest him.

[30] This phenomenon also occurs in the work of incompetent picture-makers, but is acknowledged to be deliberate in Cézanne's case.

Figure 4.28.
Georges Braque: Soda, Oil on canvas 1911.
From **Hughes 1980 p33 Fig 16.**

The work of the Cubists does not tend (as Willats claims it does) towards the elimination of viewpoint, but rather towards its detailed exploration. In the Cubists' case it is principally concerned with using painting to make a record of observations over time (Figure 4.28). They are almost non-pictures, since, unlike the Cézanne, they are even less automorphic representations of a scene then the futurist paintings which offered multiple instance in a single canvas: they are shifting from the pictorial to the schematic. It is important to recognise however that automorphism is a matter of degree. Even Cubist paintings do not require the viewer to refer to some external legend in order to decode them. The 'legend' here lies in the shared understanding of pictorial representation itself, without which these paintings would be incomprehensible. Again a spatiality has been developed which is finely adjusted to the intentions of the picture-makers and the prior knowledge of the users.

PI-realism limited to the imitation of retinal stimulation

The marks which appear in a PI-realist image must by definition imitate the source of the sheaf of rays which is received at the eye. Gregory (**1970 p33**) points out that nevertheless such images fail to specify unambiguously the shape and position of objects. Most of the ways of dealing with this problem which I have documented so far involve the use of alternative projection systems or some other means of including aspects of the scene which would not be visible in a PI-realist image, in order to provide greater clarity of information. They are interventions in ***V***, the view. However, these did not necessarily imply the *addition* of marks not authorised by the scene, the *illicit marks* to which I have several times referred, which clearly belong to ***P***, the depiction itself.

When Gell (**1998 p165**) says as part of an extended argument about the differences between language and visual art that whereas the 'd' in 'dog' does not stand for part of a dog, any line which is included in a drawing of a dog does represent some part of a dog, he is only partly right. It is true that the automorphism of pictures guarantees that the mark appears in approximately that place in the picture which corresponds

Figure 4.29
Jane Walmsley: postcard of a woman in the costume of the 1840s
© J. Bartholomew 1988

to the location of that part of the animal in the 'sheaf of light rays', but the line is not necessarily a delineation of some piece of the referent. The line may also, or instead, function as an affordance of an analogous visual experience. This is a fundamental difference between PI-realist and VE-realist picture-making.

As indicated earlier, the commonest case of the illicit mark is the drawn outline. These are rarely a crude delineation of the edges of objects. Even in Figure 4.29, where the delineation of objects is mechanical in character, a heavier outline is used around the figure to 'lift' it off the page. Though an analogue of depth perception, it of course has no direct optical equivalent in the scene. It is used to *construct* the space, however coarsely.

A related technique is used in some cartoon drawings, where a shape outlined in black is additionally separated from the background by a white outline (**Kurlander, Skelly and Salesin 1996 p229**, Figure 4.30). It is also used in graphical user interfaces to make the cursor 'float' over all other displayed elements.

Figure 4.30
The use of an 'illicit' halo to help characters stand out from their background.
From **Kurlander, Skelly and Salesin 1996**

Such illicit marks have a long history. Deregowski (**1984 p42**) shows Bushman petroglyphs (not illustrated) from South Africa, of 'ancient but unspecified date' in which the more distant horn of antelope is clearly drawn as separate from the head of the depicted animals. The illicit gap is there to 'make' space, not to match the optics of the scene.

A common technique for creating depth in drawings is to draw lines 'onto' surfaces, for example tracing round a cylindrical form or across a rectangle. Such marks may indicate the direction of surfaces converging on (or in some other way relating to) a vanishing point (Figures 4.31 to 4.33). The way some marks trace around a surface is analogous to touch (Figures 4.33 to 4.35). In neither case do they correspond to traces in the optical image. They make a space which is analogous to the perceived scene.

Such illicit marks are clearly conventional: though they build upon the basic depth cue of linear perspective, the viewer should interpret them as codes to show depth, not as actual marks on the surface of objects. Illicit marks are dependent on shared culture which enables the viewer to interpret the marks in the correct way and in so doing to see 'through' the codes without being aware of them as such.

Figure 4.31.
From a letter from Evelyn Dunbar to Charles Mahoney, September 1935 © Estate of Evelyn Dunbar.
From Fine Art Society catalogue of exhibition 'Charles Mahoney 1903-1968' at Fine Arts Society, London 10 March -14 April 2000, p13 Fig.R.

In Figure 4.31 the angle of the marks in the various cross-hatched shadings conveys the depth-spatial direction of surfaces, most notably for the changing angles where surfaces of the topiary bushes meet.

Figure 4.32.
Agnes Miller-Parker: Woodcut illustration for HE Bates 1979
Down the River Victor Gollancz, London (reissue of publication of 1937), p139.

In Figure 4.32 the angles in depth of the surfaces of the rock are created by illicit use of cross-hatching, mainly executed using marks parallel to the edges of rectangular surfaces.

Figure 4.33.
François Boucher: Young man sitting, Study of Head and Hands (detail)
From Royal Academy of Arts / Metropolitan Museum of Art catalogue for exhibition Chardin, Royal Academy of Arts 11 March - 29 May 2000, p93 Fig.13.

The lines drawn 'around' the surface of the face in Figure 4.33 have no basis in the optical image. They illicitly 'borrow' cues of linear perspective and are also rather like the traces of an exploring touch.

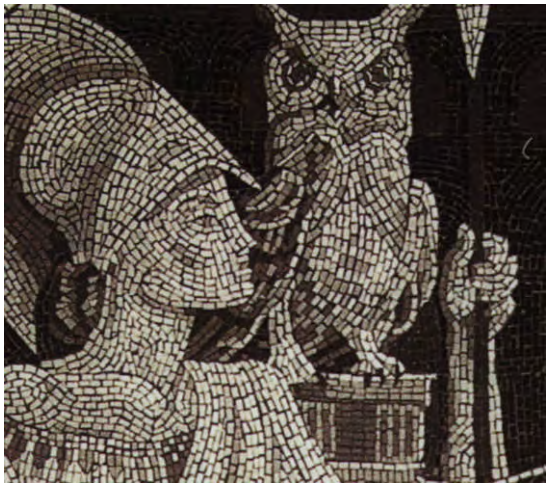


Figure 4.34.
Marjorie Knowles: Ceramic mosaic *Minerva* (detail).
From Cheek, Martin 1998 *Design Sourcebook: Mosaics* New Holland, London, p87.

Much of the work of constructing the depth-spatial form of the figures in Figure 4.34 is done by the linearity of the mosaic tesserae which travel across and along the principal directions of surfaces.

Figure 4.35.
Van Gogh: Boy with Sickle, crouching, black chalk and watercolour, October 1891 (detail).
From Harris, Nathaniel 1982 *The Art of Van Gogh* The Hamlyn Publishing Group, London, for WH Smith, p20.

The legitimate linear cues provided by the creases in the back of the boy's waistcoat in Figure 4.35 are supplemented by many other linear marks which help to 'explain' the shape of the surface. The viewer tries to decipher the marks as evidence of the manual, perceptual and constructive processes which produced it, 'reading' the

intentions of the picture-maker. Note also the exaggerated shadow on either side of the nearer arm which helps to 'make' space between it and the body behind.

Picture-makers do not adopt a rigorous logic in their use of illicit marks: such marks are usually combined with optical data in an *ad hoc* pragmatic way, which as noted earlier may be driven as much by the ongoing solution of pictorial problems as by any preconceived system. This also is something that the acculturated viewer is able to interpret. Wollheim refers to the picture-maker building up *analogies* between the medium and the object of representation, seeking an 'ever more intimate *rapport* between the two experiences' (**Wollheim 1980 p224**) and Podro offers specific examples:

> ...we recognise a figure in the lines of a drawing and when it is a figure in movement we may recruit the apparent impulse of the line – imagine the impulse of the line – to fill out our sense of the movement [...] in such cases the line relates itself to the figure twice over, once by its shape and once by its apparent impulse. The line connects shape to movement as they can be connected only in drawing. Shape and movement become projected onto each other, so that while making recognition more replete the image takes on a structure which has no equivalent outside depiction.
>
> **Podro 1998 p9**

Only in the case of the PI-realist representation can the model ***M*** and the transformations of viewing and picturing ***V*** and ***P*** be distinctly separated. As soon as one departs from strict PI-realism the two begin to interact – the *how* begins to become the *what*.

In the digital image the development of rendering algorithms has largely been devoted to resemblance to an idealised version of photography and little has been done to explore alternative relations between the evidence in the scene and the marks in the image. Lansdown and Schofield (**1995**) however list some examples, including Schofield's own *Piranesi* renderer. Attempts have been made in both 3D rendering and paint systems to replicate the materials traditionally used by artists, but the authors point out that this is in its own way a limiting view. What is required, they argue, is not just the facility to imitate familiar rendering styles, but techniques to make more expressive representations. These would be drawings which 'speak' to the viewer more directly of the forms they depict – mark-making as a form of explanation, even though it maintains much of the automorphism of 'pure' pictures. I have shown how almost all pictures other than photographs incorporate additional marks, or suppress optical information from the scene, with the pragmatic intent of articulating the space (and other aspects of the model) rather than simply re-presenting optical data. Illicit marks are a clear illustration that both ***V*** and ***P*** are representational interventions, not straightforward translations. They make it possible for the picture-maker to 'tell' about the scene graphically rather than simply to 'show' it, favour the affordance of certain perceptions over others. To create only PI-realist pictures is to forego this possibility.

8 Maturity

An aspect hardly discussed so far but implied throughout, is that picture-making has become a *mature* form, in which spatial practices have evolved to allow a close match between the particular spatialities adopted for an image and the objectives which it serves. Renaissance linear perspective offers an instructive example of transition from a problematic new technology at odds with the pictorial practices which it invaded to a thoroughly integrated informational and affective device. Greenaway has complained of the damage caused by the move to using size perspective in place of semantic sizing based on significance: 'All those other things that the Renaissance taught us to forget – that Christ is this big [spreads arms] and the apostle is this big [much smaller], which a twelfth century Amiens peasant wouldn't have had a problem with, but subsequent to the Renaissance we all have a problem with because the Renaissance taught us about illusionism and realism and all those other irrelevant phenomena' (interviewed in **Melia and Woods 1998 p30**). Despite Greenaway's objections, geometric perspective has been so absorbed into spatial practice that a 'double game' can be played in which the 'accidents' of perspective can be used as part of the articulation of meaning. This is apparent in Ucello's *Battle* (Figure 3.08) in the way that the orthogonals point at the empty space into which the Duke's horse is about to leap – a more dramatic composition than if the orthogonals had simply directed the eye straight to the Duke *in situ*. An example of both the problems and the potential is offered by the Crivelli *Annunciation* (Figure 3.07)**.** On the one hand the dictates of 3-spatial coherence have demanded the absurd introduction of an arched aperture in order that the divine ray can pass unobstructed from the heavens to the Virgin. On the other hand there is a subtle use of orthogonals as two-dimensional devices, for example to connect the distant gazing man who shields his eyes to the aperture just described. Speaking of the *Rape of Helen by Paris* by a follower of Fra Angelico, c.1450 (not illustrated), Bann suggests that the painting presents a kind of redundancy, in which those elements which are most important in the narrative *are at the same time* most important to the construction of the space (**Bann 1987 p87-9**). Importance, which had been represented schematically, came increasingly to be represented perspectively. Linear and size perspectives which had originally been difficult to put to use and risked undermining the expressivity of the medium, became part of a mature, integrated spatial articulation. Instead of the space operating at odds with the meanings of the painting it comes to enhance or even multiply the meanings. It also enables new meanings to be created which were not capable of being uttered previously, as for example in the much discussed perspectival play of Velasquez in *Las Meninas* of 1656 (**Searle 1980 p247-258** and **Foucault 1974** (**1966**) **p3-16**). Such maturity will be seen in the next chapter to be an important characteristic of the spatial character of film and to a large extent of television. However I shall argue that no such maturity has yet emerged in the spatiality of digital interactive media.

In this chapter I have analysed the notion of realism in spatial terms. I have shown that the idea of realism is often invoked without clear definition and as though it required none. In place of this vague concept I have proposed that many different kinds of realism can be discerned in pictorial artefacts, often arising directly from the selective employment of the depth cues which were discussed in the previous chapter.

I have defined one theoretical realism, PI-realism, which would give the illusion that a picture was not a picture but a view of a real scene. By definition such a picture need not take account of the cultural context, only of universal characteristics of the physiology of the human eye. I defended this concept of PI-realism against relativist views of picture-making which question the special superiority of one projection system by claiming that it has a wholly cultural basis.

I then discussed the many difficulties which PI-realist depiction raises. These included practical difficulties of execution and of controlling the conditions of reception. I also discussed several ways in which pictures which attempt to conform to PI-realism may be less expressive than if they aimed to capture aspects of visual experience more broadly conceived, which I termed VE-realism. I showed that this failure of expressivity may involve both information and affect: the PI-realist image may not function well as a vehicle for spatial information, and it may fail to afford the user the particular relation to the image which is wanted. It also cannot reflect the purposes of the image considered as part of a process and (as sketches show) excess of information may limit expressivity. PI-realism is in one sense a definitive theoretic realism, but also a limited interpretation of the concept.

VE-realism might be regarded as a subset of PI-realism on the grounds that it falls short of offering an illusion of looking at a scene. However, the arguments of this chapter suggest that they are better conceived as overlapping approaches (Figure 4.36) since VE-realism is able to capture aspects of experiencing volumetric space which are not accessible to PI-realism. It includes aspects such as non-optical projections and illicit marks which would be excluded from the PI-realist image.

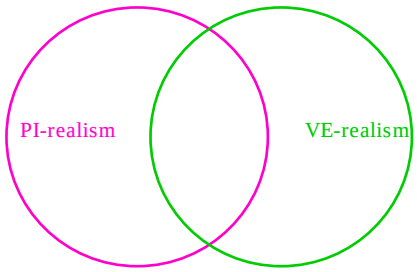


Figure 4.36 Neither PI-realism nor VE-realism should be conceived as a subset of the other.

In many cases PI-realism and VE-realism are incompatible: in order to better present some aspect of visual experience, an aspect of PI-realism is abandoned.

A true PI-realist image would make no use of pictorial convention (it would not need to). In trading PI-realism for VE-realism, the freedom from cultural convention which PI-realism offers is sacrificed. VE-realist images work to a considerable extent by being natural to the extent of automorphism, but also partly by *appearing* natural even when they are actually using conventions. As a result, to one observer a particular VE-realist picture may appear the exact counterpart of natural vision while to another it may seem strongly conventionalised. Different cultures at different times, and to a certain extent different individuals, will either see through, or be conscious of, these encodings. When Klee or Gregory look at an image having convergent verticals, they are struck by its oddness; they are looking at a picture which *seems* aberrant from natural vision. This apparent aberration is admitted by both to be in some sense correct, but the damage has been done: because both are accustomed to verticals being 'corrected' in both painting and photography, it seems to them that an unnatural encoding has taken place. This is the key to 'realism' in depiction. It is not the match of the depiction to the scene which matters, measured in objective terms, but the perceived naturalness of apprehension. It is defined not so much by the presence of any particular qualities, but the absence of one, namely by the *absence of any apparent encoding*. It matters little whether some aspect of a depiction operates through visual encoding (apprehended with the assistance of cultural convention) or through some kind of direct mapping of the scene to the plane, provided that the methods of the picture and the culturally influenced expectations of the viewer together produce this absence. It is the sensation that encoding is not present, not the *actual* absence of convention, which creates the impression of a natural, realistic picture. The following chapters will show that whether or not a depiction seems transparent in this way also differentiates spatial practices in screen-based media.

I have developed a model of picture-making as pragmatic in two senses: a picture is situated in a context of intentional activity and its appearance is moulded to its objectives; in addition picture-making is characterised by pragmatic pictorial decisions in which the elicitation of a response is as important (often more important) than the matching of an external referent. The marks which constitute a realist picture are best thought of as stimuli intended by the picture-maker to *afford an experience analogous to* natural vision, rather than necessarily being a copy of anything. The picture-maker offers data which, for the viewer informed by prior and current context, afford the 'perception' of more than is on the picture-plane.

In this chapter the difficulty of matching static pictures to the dynamic experience of natural vision has been highlighted several times. Film seems an obvious solution to this problem. However in the next chapter where filmic and related spaces are analysed it will once again turn out that the concept of making rather than matching is the key to understanding the spatiality of these media.

5 Screen Space I: the Spatiality of Film

1 Introduction

This section of the thesis, comprising three chapters on screen space, begins with an analysis of the spatiality of mainstream fiction film. The contrasting spatial practices of factual television are then discussed, forming the groundwork for a study of the spaces of interactive pictorial media.

I argue that film illustrates very clearly how the expressivity of a medium, as with pictures, is based on the ways in which spatial representations do *not* match vision, once again questioning any simple notion of realism. As before, this expressivity has two aspects – the ability to convey meaningful information and the ability to promote a particular relationship between the viewer and the representation. By understanding how the screen genre of classical fiction film makes characteristic use of space, one can begin to see which aspects of film can – and which cannot – be adapted to other genres.

There is a danger that I may seem to treat film as though it were a matter of formal design – as though the elements of sound, of narrative, and of drama were secondary. However the aspects of narrative and drama are crucial to the argument. Bordwell suggests (**1985 p50**) that 'Hollywood cinema subordinates space to narrative causality' and I shall argue that this relation between spatiality and narrative in film is so fundamental that it makes difficult any simple re-application to interactive (that is, less strongly narrative) media.

Outline of the chapter

The chapter begins with an analysis of the components of filmic space which is broadly divided into the space depicted within shots and that created between shots. Spatial properties within the frame are then considered in some detail, emphasising as in previous chapters the variables of viewing and picturing. To this is now added variation over time. Whereas one might have expected that the addition of motion and other time-based properties such as dynamic focus to depiction would lead to closer imitation of natural vision, I show that it is in fact used quite differently.

Though I make clear that the picture within the film frame does not straightforwardly capture an objective optical truth, it is clearly in some broad sense realistic because of the high level of automorphism which photography offers. Consideration in spatial terms of the relationship *between* shots reveals a quite different situation, and much of the chapter is devoted to analysing this aspect. I show how film-making, like the making of still pictures, adopts and rejects kinds of realism on the basis of its objectives. For the fiction film this is, above all, narrative.

Returning to the model of depiction proposed earlier, I demonstrate the mutual relationship of viewing ***V*** and picturing ***P*** to the model ***M***, in particular rejecting any view of film-making as the capturing of pre-formed scenes. I discuss the principal

influences on shot selection and introduce the concept of the *optimal view*, that shot which, crudely speaking, gives in context the greatest information, and which for that reason is the shot which the viewer most wants (or needs) to see. This simple concept is refined and its limitations considered. In particular I emphasise the importance of authorial denial of the optimal view for narrative purposes.

Gance remarked of filmic innovation that 'what now appears the simplest of things may once have seemed the most incredible of inventions' (**Brownlow 1968 p528**) and the hundred year history of film is filled with examples of both sudden and subtle refinements to existing spatial practice. I consider how the spatiality of fiction narrative has diverged from that of non-fiction screen genres using the example of two historical films, *Citizen Kane* and *Napoléon* to emphasise spatial techniques which have at one time flourished but subsequently been expunged from mainstream film-making. I show how this follows from the fiction film's objective of appearing unmediated despite the cavalier authorial way in which space is treated. These abandoned spatial techniques are shown later to have found their place in some specialist film-making and in factual television, suggesting that it is *genre*, rather than media or technology, to which spatial practices belong.

Though I emphasise how analysis reveals the artificiality of film, it is essential also to deal with the powerful sense of directness, of 'just seeing', which film promotes. Harrington points out (**1973 p10**) that film shows things rather than explaining them. Reeves and Nass claim (**1998**) that people react to on-screen stimuli such as a face in the same way that they react to actual faces. Sudden movements or vertiginous scenes produce a visceral response. Much of what film shows has this direct, apparently unmediated quality and it would be a poor theory of spatiality which failed to deal with this sense of the real. The issue of film realism is addressed in the same terms as for pictures in the preceding chapters, considering both *correspondence* to aspects of natural vision and the *functional* arguments against attempting to imitate vision. As before I argue that, even if it were possible to confine the design of representations to the imitation of natural vision, this would involve a disastrous loss of expressiveness. Nevertheless the *illusion* of unmediated vision is of the greatest importance.

To point up the spatial practices of mainstream film-making and to introduce the contrasted spatial practices of factual television, I discuss the contrasting spatiality of the films of Greenaway, showing how this too is an outcome of the objectives of the artefact, objectives quite at odds with the mainstream. This close fit of spatial practice to objectives is argued to be an aspect of film's *maturity*: what film-making attempts to convey and the effects that it attempts to exercise on the audience are well served by its spatial forms. This contrasts strongly with pictorial interactive media which have yet to find spatial forms appropriate to their objectives, partly because in most cases those objectives are unclear. Interactive media also suffer from the characteristic problem of any immature medium that there is a lack of shared knowledge of a body of conventions between makers and users of each multimedia genre. I show how

film's spatial codes by contrast rely on substantial prior knowledge shared between film-maker and viewer, an advantage not yet possessed by less familiar genres. Apparent limitations have been turned to strengths: the fact that the camera is not an eye allows it to do things the eye cannot do (just as in the previous chapter the fact that a picture was not a replica of natural vision allowed many possibilities to be exploited).

An indication of maturity is the degree of omission made possible by the extent of the knowledge shared between makers and users. Film has been gradually paring away the structural devices it uses, so that the simple cut is now dominant, its interpretation relying on what precedes and follows the cut and on the shared understanding of film 'language'. While it may be attractive to produce a tidy theory that structure *p* means *x* and that whenever *y* is meant it should be indicated by *q*, in practice spatial devices will always be more prone to 'messiness' than such theories allow. This 'messiness' arises from the pragmatism of spatial development, in which innovations become absorbed into general practice if they seem to work, regardless of whether they fit a neat theoretical model. I emphasise again the role of *ad hoc* solutions created during the execution of works, solutions which may be abandoned or absorbed into general practice and which take the film-maker into territory both more rewarding and more problematic than attempts at 'straightforward' representation – and I continue to show that representation is never straightforward!

Terms used

I use the term *film-maker* in order deliberately to leave open whether the decisions are taken by the director, the cinematographer, the editor or someone else. As with the previous terms *designer* and *picture-maker*, I do not suggest that one person is solely responsible. In film-making it is more likely than not that the finished artefact is the product of decisions by many individuals.

The phrase 'classical Hollywood cinema' was coined by Bordwell, Staiger and Thompson (**1985**) for their exploration of the relationship between the form of film and the context of its making. Here I use 'classical fiction film' as shorthand for a collection of filmic practices associated with mainstream commercial cinema.

2 Film space

I showed in the last chapter that what seems a naturalistic representation in one context seems 'encoded' in another. This arose partly because of the difficulties of achieving a definitive visual realism and partly because such realism was at odds with many other objectives of making pictures including the offering of other, incompatible kinds of realism.

The high level of spatial naturalism which photography gives to the individual film frame is *not* carried over into the spatial qualities of film considered over time. This is at first surprising, since it was apparent when considering pictures that a key factor militating against the illusionistic imitation of vision was the absence of motion. One

might expect that with the reincorporation of the missing time component most of these 'problems' would be solved. This turns out not to be the case.

Obviously motion confers practical benefits. It is no longer necessary to fix the viewer to a single viewing position in order to prevent the breaking of certain pictorial illusions: the impression of solidity can now be generated by moving the camera viewpoint around the model. Nevertheless there is still a problem of visual illusion being undermined by the intrusion of unwanted context. This is traditionally partly solved in the cinema by projection in a darkened room.[1] Viewers virtually reorient themselves to the correct viewing station to an extent sufficient for a general sensation of visual realism, even though there is no possibility of being deluded that the screen is a window on a real scene.

The components of film space

The essential characteristics of filmic spatiality which differentiate it from previous technologies of the moving image are (1) its use of photography, with all that that implies about both actual visual realism and perceived authority, and (2) the use of movement orthogonal to the picture plane. Previous technologies had allowed movement parallel to the picture plane but it was orthogonal movement that made the Lumières' arriving train so startling (**Brownlow 1968 p4**).

In discussing pictures I noted the difference between the extent of the model ***M*** and that of the view ***V***, which in film is one of the principal means of articulation. It is also of course one of the ways in which even the most realistically inclined film fails to be like natural vision, replacing the gradual transition from focused to peripheral vision with a strictly bounded view which is normally narrower than the natural field of vision.

The film-maker controls the actual physical spaces chosen or constructed (that is, the model ***M***), the disposition and motion of people and movable objects within the space (the action, also an aspect of ***M***), and the view ***V***. Though the photographic basis of film denies the filmmaker some of the control over picturing (***P***) which painting allows, it will be seen that even photography allows a surprising degree of control over how a view is rendered, and the use of these variables yields specifically spatial outcomes.

Movement of characters and objects in and around the space has the same significance it would have in real life, but in addition the movement takes place relative to the viewer's line of sight which alters how it is interpreted. Model, action and view are intimately related. Actors are positioned to afford the best view for the purpose. Sets will normally be constructed with particular shots in mind. In this sense it unwise to think of ***M***, ***V*** and ***P*** as independent, or even of ***M*** as preexisting: they are designed in relation to one another. In one sense ***V*** and ***P*** produce the only space that

[1] However for Singer (**1998 p45**) 'these conditions explicitly preclude our confusing the image with any prior reality: the flatness and two-dimensionality of the surface, the enormous size to which the objects have been magnified, the artificiality of the shimmering light in the darkened hall.'

matters, in that the resulting picture is the only visual evidence the viewer has. This is quite unlike a virtual environment in which the user might choose to look anywhere: the film-viewer may look only where the film-maker allows.

Spaces within spaces

The views seen by the film-viewer have a relation to the larger spaces of which they are a part (Figure 5.01). The *diegetic* space ***a*** is the inferred total space in which the action is understood to occur, including the space beyond doors which are never opened and round corners which are never explored. The viewer does not expect to see every part of this space.

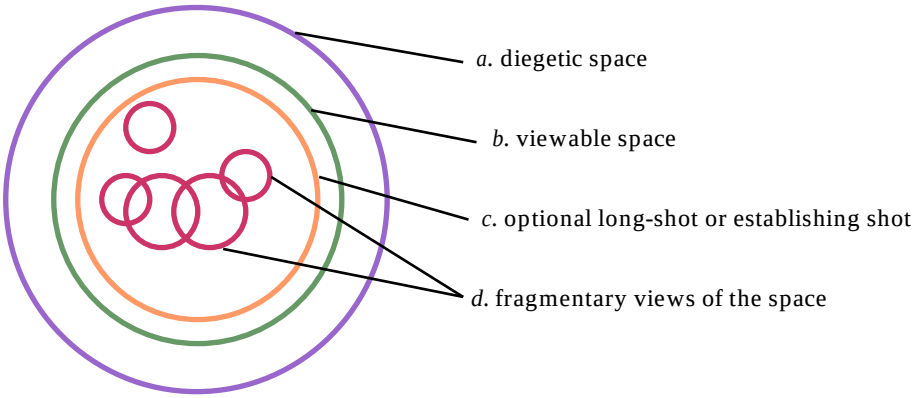


Figure 5.01. Schematic representation of the spaces within spaces in film.

A subset of the diegetic space is the viewable space ***b*** which the viewer can infer on the basis of the shots shown. Viewers will generally have the impression that they know what the viewable space looks like, even when they may not have seen it all. The shots of the viewable space may be very localised, such as close-up views ***d***, or may include a long or establishing shot ***c***.

Viewers believe that the images they see are views onto a larger spatial whole even though the views may never amount to more than a fragmentary presentation. They infer the internal spaces *between* the viewed portions and also believe in the existence of both a viewable and larger unviewable diegetic world which extend *beyond* the boundaries of the sum of all the views. This ability is assisted in film viewing by the resemblance of photographic pictures to the real world – the automorphism *within* pictures 'spills over' into an assumed automorphism of the relation between shots – so that the viewer brings to it an understanding of how the real world works and looks. In addition the viewer is greatly assisted by convention: it is partly practice in film-viewing that enables the viewer to assemble discrete shots into a coherent space.[2] These points are further developed later in the chapter.

3 Variables of viewing and picturing

Films are normally considered to comprise a hierarchical time-wise structure in which the smallest unit is the *frame*. A *shot* consists of frames in uninterrupted sequence; a *scene* is a series of shots that the viewer understands to be taken at the

[2] However Hochberg (**1987 p604**) suggests that this integrative process poses challenges for theories of visual perception conventionally rooted in vision of the natural world.

same location during a particular period of time; scenes in turn comprise *sequences*, a series of scenes related in location, time, generating action, point of view, or cast (**Harrington 1973 p8-19**). Since the viewer is not conscious of individual frames, the shot is the minimal significant formal unit in time.

Though the contents of the frame may broadly map Gibson's 'sheaf of rays,' the film considered as a whole does not, and even within the frame many techniques are used to control just *how* the image maps the scene. Each of these affects the space constructed. For Thrift (**1996 p279**) video and film are 'still essentially mimetic media that correspond to the optical wavelength of the spectrum.' This remark belongs to the school of thought that regards photography as wholly 'automatic drawing' as though the photographer or cinematographer were somehow powerless to affect the result. Though the automatic and effortless capture of appearances is one of the merits of photography for the film-maker, there are many ways in which the cinematographer can intervene to control the relationship of the shot to the natural scene, even before the use of more than one shot is taken into account. These interventions involve controlling the variables of the view ***V*** and how it is rendered in pictures ***P***.

Such variables include the viewing position and the target of the view (and thus the angle between the two), movement of the view in relation to the scene, characteristics of the lens (especially the angle of view), focus, the use of effects such as superimposition, lighting (arguably an aspect of the model rather than the view, but designed entirely with the specifics of viewing in mind) and the choice of filters, film stock and processing. All but film stock and processing may be altered continuously within a shot, so most of these variables can change with time.

View and viewing angle

Basic variables are the decision from where and towards what point each shot is aimed. Perspective painting and drawing imply an original observer (though I have discussed how in non-photographic picture-making the image presented to the viewer may not be geometrically consistent with a single viewpoint), and in the case of lens-based film-making, this original observer really exists, in the form of the camera. The view presented to the user in any given shot is definitely a view from one particular place. The same applies to views computed for a synthetic camera in a three-dimensional digital environment.

Even when a shot does not itself alter over time, it still takes place in a context of time and this is enough to make a significant difference. A series of suitable fragmentary shots will, in the absence of contra-indications, be seen as representing different aspects of a coherent space: angle of shot is a key means of achieving this. For example a view upwards to a person in a high window may be followed by a view through a window down into a garden: the spatial relation between them is constructed by the viewer on the basis of the coherence of these angles. Shot angle here is used principally as an informational device. In addition, shot angle has a

relation to the film-viewer: it is this for example which causes an upward view of a character to imbue that character with authority. This is the affective aspect. As Harrington puts it (**1973 p77**) a film-maker 'tells the viewer how to feel about a character or an action by a shot angle,' one of many examples of how film's authorial, narrative character dictates its spatiality.

Generally there is an expectation by the film-viewer of a certain 'normality' to each variable of the shot, and the further the deviation from that norm the greater is the sense of individualised feeling and attitude. This may be interpreted as a subjective quality arising from the disposition of one or more of the characters ('subjective camera' is discussed below) or as part of the feel of the film. For example, a view which exaggerates the steepness of a flight of steps may suggest a subjective view by a character, or may be part of the film-maker's own conception. It should be noted that there is an unestablishable dividing line between making a flight of steps look steeper than it 'really is' and shooting it so as not to accidentally diminish its natural steepness. I noted in the last chapter the problem of subjectivity in relation to the heights of distant hills and pointed out that a view which makes the mountains look very small – as they 'really are' in optical terms – is not necessarily a good match for subjective visual experience.

The film-viewer must know how to discern visual subjectivity. Reed repeatedly uses oddly angled shots in *The Third Man* (**Reed 1949**) as a kind of expressionistic analogue of the distorted values of the world depicted (Figure 5.02a-c). The film-viewer understands that this does not represent the subjective impression of any of the characters (nor that the buildings have themselves been built askew). Convention is vital here, as is the inference of the film-maker's intention – what Hochberg (**1987 p608**), following the ideas of Searle (**1969**), refers to as 'the viewer's legitimate assumption that the film sequence was created with a coherent narrative or expositional purpose'.

Figure 5.02a-c. Carol Reed: *The Third Man*, 1949, shots at 0h 07m; 0h 20m; 0h 56m

Other aspects of angle of view are the use or avoidance of symmetry and the tendency to prioritise the centre of the screen. *Gone with the Wind* (**Fleming *et al* 1939**), though it has one or two unusual spatial features of its own, can stand as typical of traditional commercial film practice. Bordwell suggests (**1985 p50-1**) that Hollywood prioritises the centre of the screen while avoiding symmetry, and certainly this film conforms to type in that respect. The centralising tendency in film is a naturalistic trait, because it keeps the viewer's attention away from the artificial frame of the image. It imitates the way in which in natural vision the subject of interest is without

fail located in the centre of the field of view. There are no scenes in *Gone with the Wind* in which any significant action or aspect of a situation occurs outside the central area.

It may seem surprising to characterise the spatiality of the classical fiction film as 'naturalistic' given Hollywood's reputation as the 'dream factory', but I distinguish, as in the last chapter, between different kinds of realism: it is other realisms, *not* whether a scene is believable as a piece of real life, which are significant for spatiality. The relationship of spatial representation to vision need not be different between making film match real life and making it match an imaginary scene. Indeed it is one of the principal objectives of the classical film's aspiration to visual naturalism that the imaginary should be made to seem real.

Asymmetry in the plane, and a corresponding tendency to organise views of moving subject matter in a three-quarter view, seem almost an obsession of the classical Hollywood film. Clearly when filming a moving troop of horsemen there are informational benefits, in that more characters can be fitted into the screen than if they were to travel across the picture plane, and there is less occlusion than if they were to travel orthogonal to the picture plane. However, the dominance of diagonal movement, typically from near-left to far-right, seems to imply additional motives. I suggest that it is an urge for naturalism, which in this case takes the form of avoiding anything which either reminds the viewer of the planar image as such – Harrington (**1973 p26**) points out that vertical and horizontal movements of characters remind the viewer of the limits of the picture's dimensions – or reminds the viewer of the conventions of the theatre which by tradition and partly for practical reasons are strongly based on symmetry.

Figure 5.03.
Victor Fleming: *Gone with the Wind*, 1939.
The approach to the Wilkes' ranch, asymmetrically viewed despite the inherent symmetry of the subject.
0h 18m

Even when the subject matter is itself highly symmetrical, as it often is in the classical building and cityscapes of the Deep South in *Gone with the Wind*, it is rarely presented symmetrically, so that even the avenue approaching the Wilkes' ranch is filmed at a slight angle, which seems almost perverse (Figure 5.03).

Bordwell remarks (***op cit* p53**) on the avoidance of orthogonals, and again *Gone with the Wind* is typical. The depth of the spaces occupied by the main characters is generally indicated by one of two methods. One technique views the set in three-quarter view so that it forms diagonals behind the character, rather like the

trajectories of the moving characters described above (Figure 5.04). The other eliminates geometric perspective as far as possible in favour of receding planes. This is further augmented in *Gone with the Wind* as in most films of its type by the use of shallow focus, so that the background becomes abstracted by the lack of resolution (Figure 5.05).

Figure 5.04.
Victor Fleming: *Gone with the Wind*, 1939.
Asymmetrically viewed backgrounds to many scenes, giving a 'naturalistic' diagonal emphasis.
0h 08m.

Figure 5.05.
Victor Fleming: *Gone with the Wind*, 1939.
Close-up of Vivien Leigh with defocused background.
0h 08m.

Bordwell suggests that such separation of planes is fundamental to Hollywood space. Part of the motive for this usage, I suggest, is the same as that for the elimination or 'flattening' of distant scenes in many of the paintings described in the last two chapters: it serves to prevent the eye being 'captured' by the spaces beyond the characters and to return attention to the foreground action (Bordwell himself remarks on the 'frontality' of the classical Hollywood space, ***op cit* p51-2**). But in addition, orthogonals and their associated deep spaces projected back from the picture plane would not serve the purpose they do in the Renaissance wall painting, partly to extend the space of the actual room. That would be counter-productive here: there is no desire to make the cinema screen seem like an extension of the cinema (even if it were possible given the sub-optimal views available to most of the audience) but rather to make the depicted space seem like an extension of the personal, possibly imaginary, vision of the individual situated *in no particular place*.[3] The criterion is whether or not a given technique succeeds in disguising the planarity of the image and the presence of technique itself.

[3] Elsaesser quotes two items of advice from the cinematic trade press, in the first of which in 1908 the cinema-owner is told to size the projected image so that figures in it are life size and in the latter in 1915 to size the picture according to the size of the auditorium. He sees this as evidence of a shift from regarding the screen as a virtual window in the cinema to seeing it as related only to the viewer (**Elsaesser 1990 p28 n24**).

Figure 5.06.
Victor Fleming: *Gone with the Wind*, 1939.
The burning of Atlanta. One of the rare examples of a full-face view of a scene, perhaps intended to act as a tableau.
1h 24m.

Very occasionally in *Gone with the Wind* a scene is presented parallel to the picture plane, such as briefly at the burning of Atlanta (Figure 5.06). There the purpose is perhaps to emphasise its iconic role as a kind of summative tableau before the film action proceeds.

The lens

The characteristics of the lens affect the overall character of the image but also specific spatial relations within it. Longer lenses offer a narrower field of view and a correspondingly larger picture of the scene, which is not equivalent to positioning the camera closer (Figure 5.07). Since in the fiction film the camera can generally be placed anywhere the film-maker chooses, lenses are seldom chosen for distance requirements but for their effect on perspective (**Harrington 1973 p55**).

Figure 5.07. Comparison of the he long lens and close-up.
An enlarged view from a distance (***a***) is not the same as a close view (***b***): both the relative distances of the objects from the lens and the degree to which objects occlude one another are different.

Lenses affect many aspects of the representation of a scene. Harrington's list of characteristics of the wide-angle lens includes: that it can provide more information than a longer lens since more objects appear on the screen; objects appear to be far apart; anything approaching the camera will appear huge and out of proportion; movement towards and away from the camera seems very rapid; as a result, action may be surprising or threatening; and though people seem more isolated and distant when seen through a wide-angle lens, they come together at a startling rate (**Harrington 1973 p66**). By contrast the long lens is more selective; compresses depth, forcing a subject against its background; eliminates all but a few key details (***ibid***). The qualities of a scene viewed with a wide lens, especially during camera movement, tend towards the sculptural while the flattening effect of the long lens

produces a more painterly aspect. Such perspectives offer trade-offs in informational expressiveness: the wide lens captures more across the scene but makes distant objects relatively small, while the long lens loses information in the plane but presents a relatively larger view of distant objects. Informational and affective aspects are both important, so that wide-angle allows more to be seen but it is also potentially 'more threatening' – it alters the relationship of the viewer to the space.

Harrington (***op cit*** **p55**) uses the concept of *naturalness* in a lens – 'an image approximating what the eye sees naturally' – corresponding to a 50mm lens on a 35 mm camera. Certainly extreme lenses produce effects which most film-viewers notice: long lenses cause extreme foreshortening which may even seem to reverse linear perspective while very wide-angle lenses cause noticeable curvature. The character of the lens may thus impinge on the 'transparency' of the representation. While a fisheye lens will probably seem abnormal to every viewer, the increasing use of long lenses may well have altered the threshold at which the viewer notices 'distortion' in this direction.

In addition to offering a rather fixed conception of 'naturalness', Harrington seems to underestimate the subjectivity of natural vision itself. For example when an observer concentrates on a detail in a real scene, this subjectively seems to enlarge detail and eliminate the surroundings (**Hochberg 1987 p608**). A long lens may imitate this phenomenon (**Sutcliffe 2000 p80-81**). Harrington also ignores the difference in the responses of natural vision to different kinds of stimuli: movement is detected in natural vision far outside the limits of focussed perception, meaning that the effective viewing angle for moving stimuli in natural vision is much wider than for static imagery. It is as though the natural visual system could subjectively support multiple focal lengths at need, a facility which the fixed boundary of the film frame fails altogether to imitate.

Zoom

Gance's *Napoleon* (**1927**) used a wider variety of lenses than had been used before in a single film, from 275mm to 20mm (**Brownlow 1983 p54**). Subsequently, the use of a great variety of lenses, and of adjustable lenses which can be zoomed from one focal length to another, has become commonplace. Zooming in on a subject serves the practical need of revealing greater detail within a small part of a scene without the cut which would be inevitable if two or more different lenses were used, but it is also an analogue of the psychological process of increasingly concentrating one's attention on part of a scene.[4] However, the relationship of camera zoom to natural vision is a problematic one and points up the difficulties which arise when an external mechanical device, whose effect the viewer may consciously notice, is used in imitation of a natural process within vision of which the viewer is normally unaware. As a result, the zoom lens, which might seem a good match for the subjective properties of natural vision, is in practice little used compared with the standard technique of accumulating impressions through discrete shots separated by cuts,

[4] Similar techniques are used in sound for film, where a sound to which viewers should attend is increased in volume relative to the background, imitating the natural ability to discern those sounds on which one is concentrating.

which one would expect to seem considerably less natural. Whether this arises because the cut-punctuated accumulation of shots is in fact a better match for subjective vision or for other reasons is a question which in the end I believe is probably undecidable, but it raises the basic question – like that asked about pictures earlier in the thesis – of what (if anything) film imitates, which I discuss below.

Again, tolerance of filmic techniques is not a constant: views which would have been found odd at one time are now used unnoticed. A technique of recent years, which may perhaps become normalised if more widely adopted, combines tracking (camera movement parallel to the line of sight) with zooming, so that the character of the perspective changes through the duration of a shot. The closing shot of *La Femme Infidèle* (**Chabrol 1969**) used zoom-in matched with track-out (**Callaghan 1972 p74**) and in *Goodfellas* (**Scorsese 1990**) the same technique is used in filming two characters who sit facing one another at a cafe table. They are seen in profile with a window behind, through which is visible a street with buildings and traffic. During the shot, though the two characters remain the same size, the street-scene behind them gradually moves nearer as they speak. Currently, while it is still unfamiliar, this technique is rather noticeable, yet increasing use might lead to its becoming 'transparent' as other techniques have done.[5] Another convention will have become apparently natural.

Focus

I discussed the problematic nature of focus in relation to pictures, highlighting the difficulty of deciding which aspects of vision are to be imitated. As with zoom just discussed, the key difficulty in film is the attempt to use a technique to imitate phenomena of natural vision whose mechanism is *not* consciously perceived. As with all the variables of the shot, there is the issue of control – the fact that it is the film-maker, not the film-viewer, who decides, and whose decisions may affect not only the way each shot is apprehended but also may dynamically affect that apprehension during the shot. Since it is so bound up with the fact that film is authored, this is clearly a characteristic which will require careful reconsideration when designing for systems which are controlled instead by the user.

Focus actually comprises two variables: that distance from the lens which is in optimal focus (the location of the *focal plane*); how much of the available depth in the scene is in focus (the *focal range* or depth of field). The film-maker can choose both focal plane and focal range, and increased technological capability has brought greater choice. At one time it was difficult to achieve full focus over a deep scene because early slow film-stock required larger lens apertures which in turn allowed only shallow depth of field. Later I explore the measures which were taken to overcome these problems, and their significance for any general understanding of filmic spatiality.

[5] The technique was also used in a film of Puccini's Tosca by the Italian national broadcaster RAI, so that the architectural background behind Baron Scarpia increasingly lowered over his image as he walked forward. My son at six years old asked 'Why is that man walking backwards?' However, this should not be taken as an indication that such a technique is necessarily less natural than others which now pass unremarked. It may be simply that it is unfamiliar.

A technique which alters in meaning depending on whether it is used selectively or throughout a film is that of 'soft focus' where no part of the image is fully focussed. It is used in a sustained way for idealisation or romanticisation, but when used selectively is associated with the subjective view through a character's eyes, for example to imply semi-consciousness. Unless very slight, soft focus will be noticed by the viewer; it must seem 'right' in the context of the situation and narrative if it is not to be taken as an error. In Harrington's words, 'focus gives a filmmaker an additional measure of *rhetorical control* over where a viewer looks' (**Harrington 1973 p66** *emphasis added*). This goes to the heart of the relationship between the maker and the viewer. The film-maker attempts to control the perceptions of the audience. Yet it will become increasingly apparent that there are constraints on what the film-maker may 'get away with,' which have to do with concepts of naturalism and reasonableness in the film-maker's decisions.

Pulling focus

Combining time and a shallow focal depth, the cinematographer can manipulate the focal plane dynamically. Commonly the intention is to transfer clear definition from one actor or significant object to another, as an analogue to the process of shifting one's attention (either deliberately or through the act of noticing), similar to the uses of zoom just discussed. Such 'focus through' or 'pulling focus' is an authorial technique for forcing the viewer to attend first to one thing, then another. Towards the resolution of *Who's Afraid of Virginia Woolf?* (**Nicholls 1966**) the camera zooms or tracks towards Burton and Taylor, closing in on their clasped hands, but the focus is then shifted to a glimpse of daylight in the world outside. Not only is the viewer affected by the obvious phenomenon that something out of focus is more difficult to discern, but psychologically it is impossible to resist having one's attention captured in this way. Cause and effect are in reverse: normally shifting one's attention in depth leads to a change in focal plane; here a change in focal plane leads to a shift in attention. The film-maker's will is irresistible: the viewer can hardly choose to ignore this authorial edict. However, generally this technique will be unnoticed by the viewer *if* it corresponds to the change in attention that the viewer would have made autonomously if the scene had been real. Under these circumstances the film-maker is making decisions as it were *on behalf of* a powerless viewer. This motive informs many other authorial choices in film space. It is crucial to the balance between giving viewers the view they want or need in their exploration of the situation and the narrative, and denying them such a view.

Superimposition

Superimposition, where two images are overlaid on screen by double exposure of the film in the camera or by post-processing, also has both static and dynamic forms. In dynamic use, it is principally used to effect a transition from one scene to another. As usual, the content of the two shots and the overall context influence the interpretation, so that the one technique has many meanings. Similarly in static use, where two scenes are superimposed for a large part of the duration of a shot, various meanings are available and are inferred by the viewer on the basis of other cues.

Since the photography on which mainstream film relies is largely tied to the depiction of the concrete, the use of double exposure has appealed to film-makers as a way of showing the inner workings of the mind. A typical use is to show at the same time someone thinking and what they are thinking about. It may be used dynamically to move from a character recollecting to the recollection itself. Historically it has also been used in a rather literary way to imply metaphorical identification, such as the head of Napoleon overlaid on that of his pet eagle in Gance's *Napoleon* (**1927**), but this is a form of spatiality which has all but disappeared from film presumably because it is so perceptibly artificial: there is no question here that the film-maker is simply helping viewers to look at what they are already motivated to seek. The decay of such devices from the fiction film is investigated more fully later.

There are obvious questions of informational expressiveness concerning superimposition. On the one hand it is possible to show two related scenes at once, but on the other two overlaid images are more difficult to discern than either is alone. In addition it is a technique with no apparent analogue in natural vision, though arguably it might correspond psychologically to the way in which one can be looking at one scene while imagining or recalling another.[6] In practice, the use of superimposition is minimal compared with time-wise juxtaposition (above all the simple cut) and has declined relative to other techniques during the history of film. Certainly the literary metaphorical style has declined compared with more naturalistic forms.

Lighting

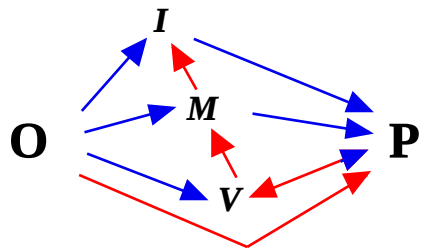
Theatrical lighting changed partly under the influence of advancing technology in the work of theatrical designers such as Gordon-Craig (**Bablet 1966** (**1962**) **p126**). The flooding of the stage with as much light as possible gave way to selective use as a way of articulating both set and action in service of the drama. Cinematic lighting has developed along similar lines, again through changes in technology as well as design practice. Psychological effects which are not specifically spatial have become attached to particular styles of lighting, from the low-key in which most of the setting is in shadow, associated with gloom and mystery, to the high-key, bright and relatively shadowless, associated with optimism and cheerfulness. High-key lighting is spatially flattening compared with the chiaroscuro of low-key lighting. All this was familiar to painters over centuries.

The *angle* at which light falls, particularly on faces, has acquired conventionalised meanings as well as having a straightforward effect on what can be perceived. Front light softens contrasts and is routinely used to make faces smoother and softer while side light has the opposite effect. Back light idealises a face if used in conjunction with front light but on its own makes it sinister. Bottom light is used for evil and top light for freshness and spirituality. The debt to the traditions of painting is obvious.

[6] Hochberg (**1987 p608**) regards superimposition as 'mere convention' on a par with the use of calendar leaves to indicate elapsed time, but clearly it is not conventional in the same way. Hochberg is assuming that film's referent is vision, but if instead it is visual imagination (as **Currie 1995 p179** suggests) then Hochberg's assertion would not hold.

Any form of lighting may change to another for dramatic effect as well as simply to alter the informational content of the image. In keeping with the naturalistic tendencies of the classical fiction film, there is generally some ostensible diegetic reason (however specious) for the lighting conditions: it will seem as though the lighting arises out of the natural characteristics of the scene, not from an authorial whim.

This relationship between the scene and the resulting picture is a clear illustration of the tight bond in film between view ***V***, picture ***P*** and model ***M*** (in presenting an idea ***I*** to serve objectives ***O***). If a filmmaker decides that some image should be dramatically lit, for example, then the set, the action and the lighting will be contrived in order to yield such an image. The process is not a pipeline in which ***M*** is designed and then subjected to ***V*** and ***P***. This 'backtracking' could be characterised as shown (in red), where the objectives demand a particular picture, which has implications not just for the view but for the model:



This puts film-making firmly in the traditions of picture-making such as painting, and contrasts with non-authorial systems where a space is designed before the user decides how to view it. This relationship of model, view and picture becomes complex in interactive pictorial media, in which different genres are beginning to adopt their own characteristic practices.

Filters, film stocks and processes

I noted in the previous chapter that the tones and colours of photography cannot be regarded as a standard of what is real, since different film stocks and processes can be used to achieve different pictures of what is before the camera. As with the filmic representation of the physical geometry of a scene, it is a moot point when the selection of any given film stock etc. makes the resulting image 'more like' nature or exaggerates some aspect of it. Similarly to the problems with differential focus, in natural vision the observer is unaware of the adjustments to the iris of the eye which allow both a bright sky and a dark landscape below it to be clearly seen, whereas unmanipulated photography favours the definition of one or the other. Filters can be used to help overcome this, making looking at the picture more like the experience of looking at the scene, even though it is a departure from the 'natural' mapping of the darks and lights of the scene which the photograph would have otherwise produced. Such attributes have a directly spatial quality when the use of a filter, for example, seems to bring a sky forward because it is more strongly defined, or a detailed, modelled surface is turned into a silhouette.

4 Space in time: shot selection and editing

I have shown that, while the photographic image imparts a high degree of automorphism to the individual shot, there is considerable scope for the film-maker to manipulate the depiction of space for narrative purposes by controlling the variables of viewing and picturing. I turn now to the spatial relationship *between* shots (which seems to depart further from realism) in order to discover the rationale for the spatiality of the fiction film. It will become clear that its objectives are decisive in determining its spatial form.

Shot selection takes place at all stages of film-making. As already noted model ***M***, view ***V*** and picture ***P*** are conceived as a whole: the design of sets is normally based on prior decisions about shots represented by the script and storyboard. Occasionally library shots are also cut into the film. The finished film is an assemblage of this material through the process of editing. Not since the very early years of the cinema has mainstream film-making consisted of pointing a camera at a scene and then showing the resulting material unedited to an audience. The time-wise juxtaposition of separate shots, intended to be interpreted as part of a single narrative or drama, became established within a few years of the Lumières' first public film-showing of 1895.

Anything which can be said about space in film can perhaps be said in terms of time – and usually is. I have been repeatedly struck by the attention paid to the non-realistic, or at least subjective, representation of time, compared to that given to similar representations of space. Because photography is broadly realistic, film as a whole is regarded similarly, despite the fact that photography has more the character of an *ingredient* in film's construction of space than itself *being* film. Lothe remarks that 'film *displays* space superbly' (**Lothe 2000 p52** *original emphasis*) but this is a very odd claim. *Shots* display space more or less adequately, but the wealth of extra-photographic innovations made over the years of film's history suggests considerable difficulties with the depiction of space: anyone who has ever made a film can vouch for the way in which simply pointing a camera at various parts of a scene in succession produces an inadequate and even misleading representation of the space.

It would be wrong to suggest that the purpose of editing is to construct and articulate space. In almost no film is the space itself the subject. The successful construction of 'just enough' space is all that is normally required. In addition there are many more constraints on shot selection and editing than the presentation of the diegetic space. I briefly enumerate these because they give clues which help answer the question of what film depicts and what the criteria are for its articulation of space.

Formal characteristics

Englander is a film practitioner rather than a professional theorist,[7] and suggests three criteria for selection: a shot should be *interesting*, *necessary* and *reinforce the dramatic situation* (**Englander and Petzold 1976 p103**). Such summary criteria raise many issues.

[7] Englander is described as having 'an illustrious career ending as Senior Film Cameraman at the BBC'.

For a shot to be interesting requires that it be interesting both in itself and in context. The latter is more significant. Interestingness is one aspect of the need to be satisfying or engaging to the viewer in a formal aesthetic way as well as in terms of narrative and drama. Hochberg (**1987 p604**) includes the provision of visual rhythms analogous to those of poetry and music as one of the functions of editing and Gessner (**1968 p263**), using the word 'orchestration', describes the formal structuring of the 'relentless stream of fluctuating lights and shapes' as one of the imperatives of film, with a view to interest and engagement, even irrespective of the story. He quotes David Lean: 'Actually I begin editing my films as I work with the writer on the script, getting a series of balances – light and dark, slow and fast, boredom and shock' and suggests that 'the aim is for effective combinations, the mixing of contrasts, differences and varieties' (**Gessner 1968 p272**). Callaghan (**1973 p77**) offers an entirely formal view of screen composition: 'when looking at a shot through the viewfinder, the cameraman must learn to interpret the action he sees in terms of the two-dimensional representation on the screen. What he is shooting is not two people running about in a meadow, but a red shape and a blue shape moving about on a flat background of green.'

Of course formal criteria and narrative objectives interact. For the relation between shots, Reisz notes a need for 'smoothness of presentation', a formal aesthetic requirement but one which is often also associated with naturalism. Complementary to smoothness is its breaking: 'a series of rapid close-ups is used: coming after an extremely long, slow-moving shot they make a striking effect' (**Reisz and Millar 1982 p54-55**). Reisz (like Englander a practitioner) recounts the decisions involved in filming five repetitions of an action, in which it is clear that the reason for filming every repetition differently is partly to create formal variety but at the same time to facilitate a gradual increase in pace.

Hochberg (**1987 p604**) emphasises the role of editing in recapturing visual attention once the content has been identified and the viewer's visual interest has waned – an interplay between the depiction and the depicted. DW Griffith was probably the first to cut actions scenes at an accelerating pace in this way. Pace in cutting is an equivalent of *illicit mark-making* in pictures, in that the pace of the shot-movement and editing is independent of the pace of the depicted action – it is not derived from its appearance – but affects how it is perceived. In any figurative film there will be an interplay between the formal patterning of film and its depiction of situations, characters and events; this is analogous to the relationship between the formal qualities of a picture and the depiction it offers discussed in the previous chapter.

'Necessary' shots: informational economy

Clearly the requirement to be *interesting* will sometimes be in conflict with the requirement for economy of expression suggested by the word *necessary*. Perhaps surprisingly, the demands of functional information efficiency are very important in determining the overall spatiality of film. In particular this relates to the concept of the *optimal view* which I describe later.

Englander's 'necessary' is useful in emphasising the pragmatic nature of film-making in which some totalising kind of 'capture' is irrelevant. Relevance is a strong influence on shot selection and hence on film's spatiality – no shot appears which does not serve a purpose. This immediately makes it unlike natural vision. The viewer's time is used as economically as possible: what in information science terms Card, Pirolli and Mackinlay (**1994**) called the 'cost of knowledge characteristic function' is just as relevant here, largely because of the need to maintain the viewer's motivation. For example, extensive filmic perusal of the setting is excluded unless it serves some specific purpose. King comments '...the set is nothing but a set. To me it is completely wrong to photograph it just because it's lavish. We are telling a story' (**Brownlow 1968 p109**). This raises another aspect of necessity: the need for every part to fit into an overall scheme of articulation, the idea of 'keeping something in reserve' so that the film is considered as a whole to which each part is contributory. For example, Callaghan offers the argument (**1973 p73**) that 'the more powerful a cinematic device, the less frequently it should be employed, especially within one film'.

Coherent space as an illusion

When shots are put together through editing, this affords spatial interpretations of the relations between them. This has a relation to error rather like the creation of depth illusion in pictures. The film-maker's task is to hoodwink the viewer into constructing a coherent space from a series of glimpses. Provided the film-maker works within certain constraints (some conventional), this 'deception' is easy to do, rather as it is easy to convey a sensation of depth in the pictorial plane using some of the many depth cues discussed previously. The film-maker offers fragmentary evidence, organised with a view to affording certain assumptions and interpretations, and the film-viewer (partly on the basis of shared conventions) duly makes those interpretations. The opportunity for error of another kind, in which the viewer infers alternative spatial relations which the film-maker did not intend, is of course always present. Hochberg points out (**1987 p606**) that most examples of what a film-maker considers bad editing seem to show movement when none is intended or conversely show no displacement where a large displacement in fact occurred. Convention can assist correct interpretation, such as in the '180º rule'.[8]

In a picture like Cotàn's *Quince, Cabbage, Melon and Cucumber* (Figure 4.09, previous chapter) the kinds of cues chosen supported, among other realisms, the realism of causal coherence so that the viewer was able to say exactly how the objects interrelated and what would happen if the objects were disturbed: the spatial relationships were *shown* by use of the cues of linear perspective, modelling and cast shadow. However between cut-punctuated shots, as distinct from within them, such direct showing is impossible: the viewer does not *see* the connectedness of the space but *infers* it. Hochberg (**1987 p607**) believes that many or even most motion picture

[8] As the film-maker cuts back and forth when showing, say, a conversation between several characters, the viewer expects all shots to be taken from one side of an imaginary line through the centre of the action. Violation of this rule leads to the wrong interpretation, namely that the characters have moved or even that they have begun talking with their backs to one another (**Harrington 1973 p131**). The 180º rule gives further clues to the pragmatic nature of viewpoint positioning. It is clear that the conventions of viewing position do not say which viewpoints may be used but only which may not (on the grounds of their leading to misapprehension). Otherwise *any* view is acceptable.

cuts occur between views that do not overlap at all, and calls on perception studies to explain the viewer's inference of coherence given that some forms of awareness which support the inference of coherence in natural vision are not available in film,[9] but he probably underestimates the benefit given by convention and by the viewer's assumption of intentional utterance on the part of the film-maker.

Though space can be treated cavalierly, it needs to be tolerably well understood by the viewer as a coherent place in order for the story to make sense: if an event is to occur within a physical space at some point in the narrative, the viewer must be able to place it within the set of spaces so far seen at the time when it occurs. In practice this often means establishing in advance the presence and location of objects which later will become significant. At best this may act as subtle premonition of an impending event, at worst as grossly obvious. In either case, it is clearly an authorial, narrative technique: if users were free to look anywhere, there is no guarantee that they would have the necessary information.

The trajectories of moving objects, such as cars or planes moving from one location to another, or actors moving from room to room, must seem to take place in a space of which – to a certain extent – the viewer has a coherent model. The realism of photography is an aid in affording coherent space between shots. If the action of a film takes place in a street, much of the film-maker's task of achieving coherence is taken care of by prior knowledge about how streets are in the world. As Currie points out (**1995 p104**) once film begins – as it does from at least DW Griffith onwards – to mix scales of view in an arbitrary way, this use of familiarity with the world is necessary even just to enable viewers to know the relative *scale* of what they see, let alone any more complex spatial relationships.[10]

With real-world subject matter, to some extent the film-maker's job is a negative one: to avoid confusing the viewer and allow the familiarity of the subject matter to do its own work of establishing coherence. This is perhaps one of the reasons why the visual realism of photography, a realism which is actual in Arnheim's sense of being a 'tracing' of the scene, is often taken to imply a similar realism, which emphatically does *not* exist in the Arnheim sense, in the film overall.

Familiarity with the kinds of space depicted also allows the film-maker to make authorial play with the perception of *time* which would otherwise be impossible. This may amount to supporting simple omission, such as that of the plane journey between two places, which derives from the viewer's knowing a variety of things about travel: the viewer's understanding of the real space which is alluded to allows the liberty with time to be accepted. But it may also enable incongruities to be created for dramatic effect, since knowing how the world is in spatial terms enables anomalous presentations of time to be detected. In Hitchcock's *Suspicion* (**1941**), intercutting between a character who stands still and one who approaches is done

[9] For example, viewers' knowledge of where their eye is directed cannot help tell how the successive views are to be put together as it can in natural vision (**Hochberg 1987 p604**).

[10] As a result, when the film-maker wishes to show objects of an unusual size such as miniature people the only method available is to insert in the scene other objects of large relative size: in Lilliput, Lilliputians are normal.

such that the approaching character seems at times to make little progress across the floor, while later he has covered a remarkable distance since he was last seen. Familiar space is used to create anomalous time: another illustration that the spatiality of film is intimately bound to its narrative character.

Other influences on filmic space

Another aspect of editing, which lies at the point of interrelation between the form and the content, the story and the plot, is its use in dramatic juxtaposition. Like many formal devices this is capable of yielding different meanings depending on the content and context. It may be humorous, playfully or painfully ironic, or take the form of a visual conceit intended simply to charm or intrigue.

These styles of editing are not intended to provide evidence on which the viewer constructs a space, but to highlight difference and disjuncture. Eisenstein's idea of montage as a collision – 'From the collision of two given factors arises a concept. [...] Montage is conflict' (**Eisenstein 1977 (1949) p37, 38**) – continues in mainstream film-making, although generally in a more naturalistic and less literary way than in his films, a development which parallels the changes in the use of superimposition commented on earlier: it is one of the naturalistic traits of the classical film that the appearance of each shot should *seem* to be internally motivated by the action, even when, as I have noted, it is in fact a gross authorial intervention for the purposes of dramatic narrative.

Since they are not intended to be seen as other parts of a coherent space such montages are irrelevant to film's spatiality in the obvious sense, but they often make use of the *planar* space of the image in order to achieve their purpose. Whereas the transition from one shot to another is by convention matched such that 'tonality, movement, and the centre of compositional interest shift enough to be distinguishable, but not enough to be disturbing' (**Bordwell 1985 p55**), in the case of montage as conflict, it is clearly some kind of disturbance which is sought. Occasionally this is done by choosing shapes which are formally similar but which depict different objects. In Peeping Tom (**Powell 1959**) coffee pouring into a cup in one shot becomes whisky pouring into a glass in the next. This draws attention to the film-maker's role, and is at odds with the modes of view used elsewhere in the film. Though I shall emphasise the naturalistic spatial tendencies of the classical fiction film, this is a reminder that no categorical assertions can be made about how film space works: only tendencies, not rules, can be discerned.

5 The concept of the Optimal View

Influences on shot selection and editing identified so far include: formal aesthetic appeal; interaction between the form and the story; informational economy; the construction of apparently coherent spaces even when these do not really exist; dramatic juxtaposition of various kinds. I look now in greater detail at one aspect of informational economy: the concept of the *optimal view*. This is a concept which helps

to explain the special character of film space and draws attention to the difficulties of transferring the spatial practices of film to non-narrative media.

In the display of textual and schematic configurations on screens, the optimal view of the configuration for any particular purpose is made available through the provision of zooming and scrolling devices or in some cases by the use of multiple panes to offer multiple views or different scales of view. In one sense the configuration is its own optimal view since it is designed in the plane in such a way as to best represent its content and internal relationships, and this planar organisation is then simply mapped to the plane of the display. In the case of pictures, as I noted in the previous chapter, the model and the view interpenetrate: how the model is shown becomes a substantive part of what the model is and optimal expressiveness is achieved through adjustments to the model, the view, and the relationship between them. In film, the tendency for the *how* to become the *what* is still more marked, since, while there may be in some sense a pre-filmic model of which the shots are pictures (albeit possibly not all in one actual unified space), the only evidence is (a) partial and (b) affected by the surrounding shots.

Editing as omission of the irrelevant

A basic criterion in editing, already noted, is the omission of the irrelevant. Film makes extensive use, partly for practical reasons such as lack of space within the frame or even the cost of sets, of synecdoche – a periscope may be all that is needed to imply the presence of a submarine (**Harrington 1973 p26**). Editing can simply omit irrelevant material: it would be ridiculous to show an entire plane journey if shots of the departure and arrival are sufficient. To what level of granularity may this principle of omission be carried? – in mainstream film practice, it seems, to any level. It is the basis of the reaction shot, for example, in which the film-viewer sees the two significant parts of a situation – the observer and the observed – in two separate shots, without the film-maker having either to show both parts of the scene in a single shot or to pan the camera from the observer to the observed and back again.

Editing as construction of the relevant

Though this characterisation of editing as *omission* is one possible model, it has two weaknesses as a description of how editing works. First, it supposes a pre-pictorial space which actually exists and from which the film-maker selects, whereas this is not necessarily the case since, as already discussed, the film-maker offers only enough evidence for the viewer to construct a space adequate to the purpose and no more. The film-maker is better thought of as *composing* a filmic space rather than *capturing* a pre-pictorial one. And the notion of omission does not give sufficient clue to the criteria for selecting those views which *are* shown. To say what decides these criteria requires the concept of the optimal view.

The optimal view: initial definition

Crudely stated, the optimal view of a scene or action is the one which provides the greatest information about the situation or event. No single viewpoint is likely to

fulfil the requirement, since a shot which conveys the overall situation may have insufficient size to show important detail while a close-up shot will reveal insufficient of the overall configuration. In comparing film to theatre, Furnham (**1999 p55**) suggests that a rationale for film's close-up shot of an individual character – perhaps the most common form of optimal view – is that it is equivalent to a component of live theatre missing from film, namely stage presence. There is no doubt that close-ups do have a special immediacy (an affective characteristic), but a more prosaic explanation is that they provide better information about the actor than would a distant view (a difficulty for which the stage actor compensates by enlarged gestures and other conventional devices). As always, history is instructive: Brownlow documents the difficulties which early audiences experienced with close-up views of faces:

> They couldn't understand how people were walking around without legs. In the theatre they were accustomed to seeing the whole body, and what it was standing on. But to see a head moving around, cut off at the neck, just wasn't acceptable.
>
> Alan Dwan interviewed 1964 (**Brownlow 1968 p98**)

Gance was ordered by an executive of his film company in 1913 not to use close-ups (**Brownlow 1968 p524**). This indicates the important difference between a technique being unproblematically realistic and therefore immediately acceptable, and its coming to seem natural through a process of acculturation.

The optimal view: modification no. 1 (context)

The information value of a shot – the characteristic which makes it *optimal* – is contextual. A shot may offer very little information when seen alone, but when inserted in a sequence may provide just the information which viewers need – what Hochberg (**1987 p607**) calls the answer, in one shot, to a 'visual question' posed in the shot preceding. In a typical fiction-film scene, when Vivian Leigh falls down stairs in *Gone with the Wind*, six camera positions are used in eleven seconds, each providing the optimal view when seen in context (Figure 5.08a-f). Aside from its affective qualities it can be regarded as the most informationally expressive articulation of the event and the reactions of the participants. Patently there is no position which could be adopted by an actual observer situated in the scene which would yield these views – an issue to which I return later.

a-b Gable and Leigh at the top of the stairs arguing: mid-shot and close-up

c Leigh falls: high-angle mid-shot

d Leigh's head on the floor: close-up

e Gable alarmed: close-up

f Gable runs down: long shot.

Figure 5.08a-f. Victor Fleming: *Gone with the Wind*, 1939. Vivien Leigh falls downstairs. Six camera positions are used in eleven seconds, each providing the optimal view of the current action in context. 3h 19m.

Whereas in some genres it is acceptable to solve the problem of showing both detail and context by displaying more than one image at once, in the fiction film it is not. This is one of the principal spatial differentiations between genres, and it will be seen how even within a single technology such as television, different genres have developed their own distinctive spatialities.[11] In the classical film, only temporal and not spatial juxtaposition of separate views is permitted.[12]

The first refinement of the definition of the optimal view, then, is that it is the view which is the most informationally expressive when seen in the context of the other views, not that it in itself is necessarily particularly information-rich.

The optimal view: modification no.2 (the psychological component)

Carroll (**1996 p125-138**) proposes a theory of point-of-view editing based primarily on the way in which humans, and indeed other mammals, acquire vital information by looking at what others look at, a behaviour with high survival value. In these terms there is a transfer to film of the everyday 'need to see' which governs observers' looking behaviour as they attempt to discern what others are thinking and intend to do. One benefit of this ecologically grounded theory is that it emphasises the aspect of need and inner compulsion: it is not a matter of idle curiosity, of *wanting* to see, but of *needing* to see.

A group of realisms which I introduced in the previous chapter concerned the empathy of the viewer with animate beings in the depiction. One such realism was basic, that of movement, and is taken care of by film's ability directly to show moving images.[13] Another was a sense of animation, of coexisting with a living being, while a third was psychological engagement, as though with another mind. This is a strong determinant of the use and articulation of film space, and adds another aspect to the notion of the optimal view. Whereas I defined it initially in general terms as the shot which conveys the most information about an action or situation, I now extend and qualify that definition to be the shot which, when the psychology of the individual or the interrelationships between characters is important (and there are few occasions in the classical fiction film when they are not) best allows the *psychological* scenario to be

[11] Of course some innovations lead to no genres and remain minority practice.

[12] I discuss later the decline in split-screen and other multi-image uses of screen space in film (and its continuance in other screen-based media).

[13] Currie (**1995 p34-42**) expends considerable effort in arguing that the movement of objects seen in film is real movement in the same sense that movement is seen in real objects, whereas others (he says) have argued that the movement is illusory, perhaps on the grounds that the film consists of a series of still pictures. For the purposes of this thesis I shall assume with Currie that the movement of on-screen objects is real movement.

understood. What makes the optimal view optimal is the contribution of such a shot, not to the understanding of the narrative *per se* (let alone the space), but to the understanding of the psychology of the depicted scene as a component of the narrative.

The optimal view: modification no.3 (the right view and the optimal view)

So far it would seem that for any given scenario or action, we could say which shots will prove optimal, and these could be selected according to a formula (indeed many lesser films are probably shot and edited in this way!). This would make fairly easy the development of algorithmic approaches to shot selection and editing. However, there are significant constraints, arising partly out of the formal characteristics discussed above, and, even more importantly, out of the fact that film is authored narrative.

Armes (**1994 p*vii***), arguing against the concepts of the screen-as-picture and the film-as-narrative, reasserts film's role in presenting and articulating drama (that is, the depicted dramatic events in ***M***). At its crudest, such a characterisation would require that the optimal view be presented at all times. This is the position taken by He, Cohen and Salesin (**1996**) in devising an automated cinematography system for use with virtual environments. However, as Armes uses Williams[14] to point out: 'The film is in one way a single recorded performance, but in another way, and more significantly, it is *in itself* the dramatic production: the actual shaping of the work' (**Armes 1994 p*vii*** *emphasis added*).

Clearly the selection of optimal views does not operate independently of the criteria discussed previously, such as formal patterning: a shot may still be selected for its contribution to the film's formal qualities as much as for its information value. I have noted how the viewer's interest may be elicited by a shifted relationship between the depiction (***V*** and ***P***) and the depicted (***M***). Armes (***op cit*** **p42**) suggests that while some action may 'seem to demand to be seen in close-up' – a phrase which strongly evokes the notion of the optimal view – only the weakest film-maker will follow such a conventional matching and that the slippage between the events and the depiction will usually in itself have significant interest for the viewer. This is the single biggest qualification of the concept of the optimal view. The selection of shots does not arise naturally out of the action: in many cases, the film-maker deliberately denies the film-viewer the very shot which would provide maximum information.

Authorial control: denial of view

In a sense all filming could be seen as a denial of the viewer's will at the expense of the maker's, since it selects a little aperture on the world, but as I indicated before in most cases the choices made are the ones the viewer would also make: in Hochberg's terms, the answer to the visual question is provided. Now I am concerned instead with the kind of denial which is experienced as such (though not necessarily consciously) by the viewer. The means of denial involve various combinations of shot

[14] Raymond Williams 1991 (new edition) Drama in Performance, Open University Press, Milton Keynes, quoted in **Armes 1994 p*vii***.

selection, the use of variables of viewing and picturing and of editing, and is fundamental to narrative media.

In shot selection a viewpoint may be chosen which in some way conceals the object of interest, such as by internal occlusion within the shot – a relationship contrived between the model and the view. In *Rosemary's Baby* (**Polanski 1968**), Ruth goes to use the phone in an adjoining room. Polanski uses the doorway as a frame for this action, so that Ruth's face is concealed from view – the viewer yearns (and expects) to see her expressions, but cannot.

The opening sequence of *Touch of Evil* (**Welles 1958**) plays several similar tricks, denying viewers the chance to keep track of the very thing they most want to see, by filming the main characters taking a journey round several blocks and down several streets, so ensuring that the view of them is repeatedly interrupted by buildings, traffic, passing hand-carts and other obstructions. The effect of the repeated occlusions is to create an atmosphere of anxiety which, as in other examples I have cited, spills over from the formal difficulties of negotiating with the medium into the perception of the story itself. It is quite unlike natural vision, in that the camera path is such that no person could possibly experience such views in normal experience. On the one hand an impression of unmediated viewing is offered by the unbroken crisply focused photographic stream of information, while on the other it is authorial in the extreme. It repeatedly denies the viewer the optimal view, in the service of narrative, and its spatiality arises out of the demands of narrative drama. Naremore calls the frustrations the camera encounters in *Citizen Kane* (**Welles 1941**), such as a door closing or a light clicking out, 'affronts to the audience's curiosity'. Their purpose is affective – to create a sense of mystery and subtle anxiety (**Naremore 1978 p71**) – but this is achieved by the denial of information.

Other variables of the shot are also used to create some form of visual denial, concealing the detail the viewer most needs to see. Some are specifically pictorial techniques such as focus, while others adapt theatrical articulations of space such as lighting. In the first category is a scene from *Lawrence of Arabia* (**Lean 1962**) in which a character approaches from a distance on horseback across the desert. The shot deliberately begins far too soon for the man to be clearly discerned, due to his small scale and the effects of dust and heat-haze. *The Third Man* (**Reed 1949**) offers an example of the second category where selective lighting – justified internally by the changing lights cast from upstairs rooms – allows at first only a view of Harry Lime's shoes.

Editing also is routinely used to deny the optimal view. At the time when the viewer is expecting to see something, the scene shifts to another aspect of the story. This may, like the variables of viewing and picturing, be disguised as having some internal cause. For example perhaps two characters pass through a door which they then close and the camera simply fails to follow them: the next shot is the start of a new scene.

Often the illusion is created that the denial of optimal view has an internal diegetic cause since in the real world doors do indeed prevent viewing, heat haze does make things difficult to identify, and deep shadows do conceal things. However since the early history of cinema, far more extreme methods, with no diegetic justification whatsoever, have been used to support narrative at the expense of naturalistic views of space. Whereas the innovative use of editing by Porter at the beginning of the twentieth century cut from one scene to another principally for the practical objective of showing what was going on simultaneously in two or more places – it had become impossible to accommodate the events he wanted in a single shot – Griffith did so with design: 'the viewpoint is changed not for physical but for *dramatic* reasons...' (**Reisz and Millar 1982 p22** *original emphasis*).

> At this point, instead of showing the assassination, Griffith interrupts the action of [shot] 36, which was probably shot as a continuous take with 38, to give a glimpse of the stage (37). [...] The view of the stage in 37 *adds nothing to our knowledge of the scene*. It is inserted for purely dramatic reasons: the suspense is artificially kept up a while longer and Lincoln's complete unawareness of Booth's presence is indirectly stressed.
>
> Reisz and Millar 1982 p23 *emphasis added*

The cutaway is inserted for purely dramatic reasons. It has nothing to do with simple depiction (which would always offer the optimal viewpoint) and everything to do with narration. Seen simply *as narration* of course this is not innovative at all – Murray (**1997 p29**) for example documents the nineteenth-century literary antecedents of filmic technique in Brontë, Dickens and Tolstoy.[15] Griffith's contribution was to see that the inherent naturalism of photography within shots could be combined with an arbitrary approach to space between shots.

Film viewers and the intentions of film-makers

The conventionalised character of point-of view, reaction shots, and other views which allow the psychology of a scene to be apprehended by the viewer raise the issue of intention. Viewers want, or more correctly (following Carroll) *need*, to see something; but in addition, viewers know on the basis of convention what shot they should be seeing at any moment. If it is denied them, they are not only denied information, but are aware that they are wilfully denied it by the film-maker. The denial is doubled: once on the grounds of information withheld and again on the grounds of expectation confounded. It is clear that the spatiality of film is not just subordinate to Bordwell's 'narrative causality' (if that were taken to mean simply

[15] There are much earlier examples: the narrative poem of the late 14th century, Sir Gawain and the Green Knight, deliberately describes its subject in a different order than any observer would see it naturally:

> His loins and limbs were so long and great
> That he was half a giant on earth, I believe,
> Yet mainly and most of all a man he seemed,
> And the handsomest of horsemen, though huge at that;
> For though at back and at breast his body was broad,
> His hips and his haunches were elegant and small,
> And perfectly proportioned were all parts of the man,
> As seen.
> Amazed at the hue of him,
> A foe with furious mien,
> Men gaped, for the giant grim
> Was coloured a gorgeous green.

Clearly the first thing the internal observers would have noticed was that there visitor was green, yet this fact is withheld until eleven lines of description have elapsed. (Sir Gawain and the Green Knight, anonymous, late 14th century 1964 translated by Brian Stone 1959 and 1964 Penguin, Harmondsworth UK) . The translation preserves the line order of the original.

considerations of comprehensible storytelling) but to many other aspects of film intimately bound up with its being not just narrative but also dramatic and psychological, and that it is experienced in the context of other films. This also will make problematic any simple transfer of its spatial practices to other media genres with different characteristics.

6 Historical divergence of fiction and non-fiction genres

I have argued that space is treated cavalierly by the film-maker in the interests of authorial narrative. However I have also argued that mainstream fiction film aspires in general to seem like natural vision. Part of the evidence for this lies in the historical elimination from this genre of spatial practices which might excessively draw attention to the fact of representation.

Classical fiction film	**Non-fiction newsreel**
all material is photographic, or in the case of matte paintings, computer graphics etc, is intended to be mistaken for photographic	drawn and other clearly planar material is used together with photography, including rostrum camera exploration of such artefacts
textual titles are avoided wherever possible, to the extent that their inclusion is nowadays often found humorous, a somewhat embarrassing intrusion from a different genre, and an admission of failure on the part of the film-	textual titles are extensively used, in some cases on a plain, non-cinematographic background
every shot is motivated, with no inclusion of material which is gratuitous	there is frequent lack of motivation for individual shots, which are presented gratuitously rather than in response to a desire or need created by the narrative
transitional effects are kept to a minimum, so that dissolves are brief, most shot changes are cuts, and no special effects are used which draw attention to the planar character of the projected image	transitional effects between shots are extensively used, including horizontal and vertical wipes, iris-open and gratuitous dissolves
optimal views are generally provided, such as close-ups of the principal characters, unless the narrative demands that they be withheld	optimal views are absent where they would normally occur (in a real newsreel this arises for practical reasons but is imitated in Welles' pastiche)
the locus of interest is centred in the frame	framing often leads the object of interest departing from the centre of the picture (again from practical causes)
any graphic material such as letters, newspapers, and so forth is situated in the diegetic space; when such material must be legible it is preceded by a view establishing its diegetic credentials and is then usually shown at an angle to imply some sort of 'natural' viewing	planar animated graphics are used
didactic graphics are never used, nor are distortions of timescale (that is representations in which the non-realism of the elapsing time is made patent)	graphics and other sequences animate in non-real-time
symmetry, frontality and the shallow space of the theatre are deprecated while oblique asymmetrical views are favoured	the symmetry, frontality and shallow space of the theatre are commonly used
generally the presence of the camera is concealed, so that hand-held camerawork is confined to the imitation of very specific phenomena such as first-person and point-of-view shots, and is even then often kept as fluid and free from vibration as possible.	hand-held camera is common where accidental movements are strongly apparent (again this accidental phenomenon is imitated by Welles)

Table 5.09. Differences between the space of the fiction and non-fiction film as exemplified by the main narrative and embedded newsreel of *Citizen Kane* (Welles 1941)

The divergent spatialities of two cinematic genres are both conveniently present in a single film, *Citizen Kane* (**Welles 1941**), which offers a pastiche newsreel within the main fictional narrative. For Bordwell (**1976 p106**) the newsreel sequence 'virtually recapitulates the technical development of cinema from 1890 to 1941.' But these developments, which had at one time been in general use in all film-making, had become characteristic of a particular genre which, fifty years into the history of the cinema, differed significantly from the spatiality of the fiction film. This allowed Welles to play 'games' exploiting the difference between the genres. The spatialities of documentary and fiction have followed divergent paths arising out of their different objectives – one to *show*, and one to seem like *seeing*. Key differences are summarised in Table 5.09.

The wide range of filmic techniques used earlier in Gance's *Napoleon* (**1927**), some of them innovative, also show techniques now abandoned in mainstream fiction film-making. Some of the film's difficulties (as we would now regard them) arise partly from its status somewhere between fiction-like narrative and documentary. Spatial practices such as split-screen and the use of non-photographic interventions in the image lie at the point of divergence between these two genres, so that what is now unacceptable in a fiction film is standard practice in modern documentary work, such as factual television. Other outmoded techniques, such as the use of literary metaphorical superimposition (extensively used in this film) have been dealt with earlier.

The demise of split-screen

Brownlow (**1968 p23**) suggests that the first use of triptych may have been by Phillips Smalley in *Suspense* of 1913 (Figure 5.10), showing three simultaneous interrelated actions at one time. Gance used a triptych for the first time in *Barberousse* of 1916.

Figure 5.10.
Possible first use of triptych in 1913 by Phillips Smalley for *Suspense.*
Taken from **Brownlow 1968** p25.

In his *Napoleon* – another film which shows many techniques now occulted from the mainstream – the hero is presented in the centre panel of a triptych (not illustrated) while the flanking panels offer contextual views of a battle. At the close of the film, in the original screening, three projectors were used in order to suddenly triple the width of the image. This involved sequences in which the three images were different shots, and one in which they formed three parts of a single panoramic shot.[16] Earlier in the film a pillow fight in Napoleon's school dormitory splits into four and then

[16] At the close of the film the flanking images were tinted red and blue to produce a tricolour. Colouring of mono-chrome imagery and any other overt interference with natural appearance has been abandoned by mainstream cinema but is fairly common in factual television.

nine panes, partly with the functional aim of showing many actions in one screen, but also presumably to add visual patterning for its own sake (Figure 5.11a-c).

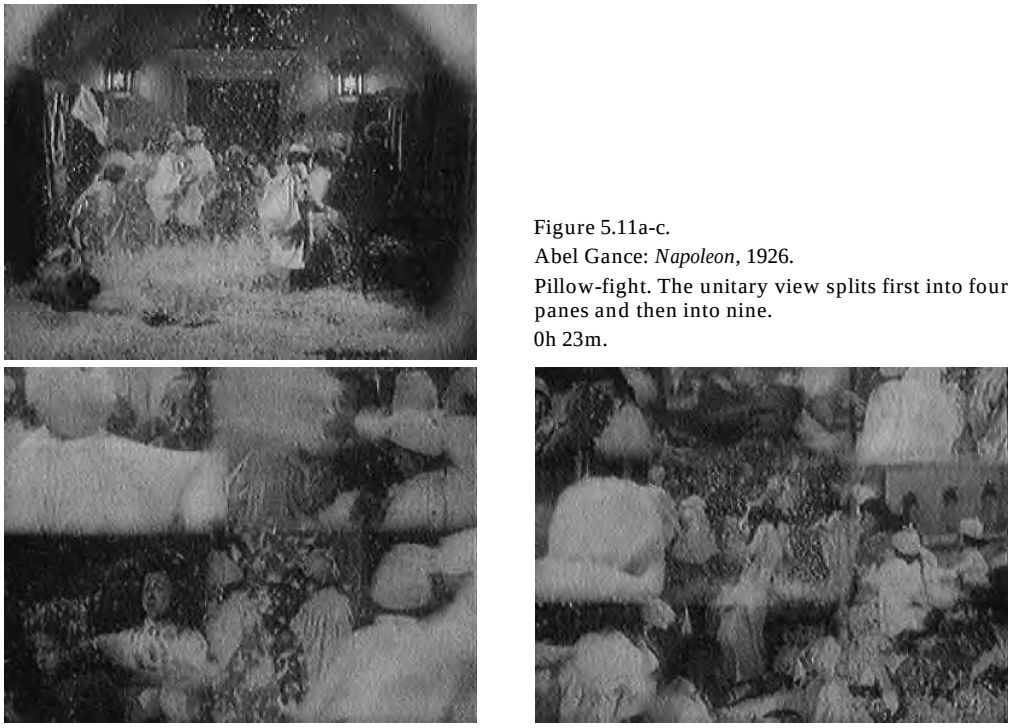


Figure 5.11a-c.
Abel Gance: *Napoleon*, 1926.
Pillow-fight. The unitary view splits first into four panes and then into nine.
0h 23m.

Split screens are not the only form of multi-pane presentation to have been abandoned by the fiction film. Salt (**1990 p32**) describes how GA Smith's *Santa Claus* of 1898 displayed parallel action using an inset image, picture-in-picture, in part of the main scene. Musser (**1991 p224**) shows the opening scene from Porter's *Life of an American Fireman* of 1902 (not illustrated) in which a scene which is being imagined by the fire chief, whom we see, is included in a circular picture-in-picture.

Modern mainstream film practice is quite different. Split screen is anathema, being generally only used in light comic films for example to show simultaneously the two parties in a telephone conversation. Significantly, the aim in such cases is to objectify the scene, drawing attention to its parts by distancing the viewer. It is not used in dramatic narratives where the viewer must be deeply engaged. By contrast, the use of multi-pane images is increasing in factual television, where 'transparency' of viewing is not a concern.

Captioning and other non-photographic interventions

Being a 'silent' film, *Napoleon* is obliged to use intertitles, but it also uses captions in other ways. At one point (Figure 5.12) a title informs the viewer that the filming on Corsica was done in the locations of the actual historical events. This would be acceptable in a modern documentary, but not within a dramatic narrative. This title is an intrusion from another mode of representation which, quite literally, draws attention to the fact that this is a film.

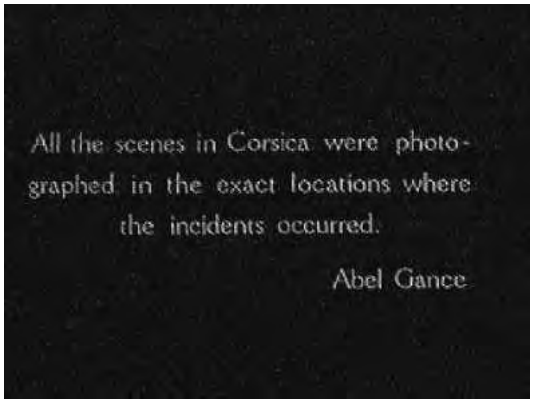


Figure 5.12. Abel Gance: *Napoleon*, 1926.
A title vouches for the authenticity of the filmed scenes.
1h 06m.

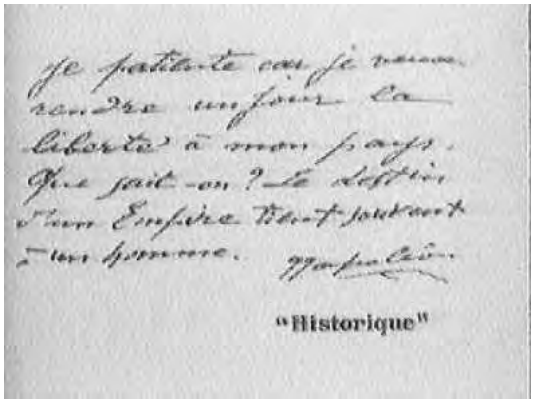


Figure 5.13. Abel Gance: *Napoleon*, 1926.
A reproduction of a letter is captioned 'Historique' to indicate that it is not fictional.
0h 18m.

Similarly, a letter written by Napoleon is shown in diegetic context – he is seen writing it – but when presented frontally so that the viewer may read it it is captioned 'Historique' to assure the viewer of its authenticity (Figure 5.13). In breaking both the spatial and representational coherence of the film this is to modern eyes counterproductive, another case where the spatial practices of historical dramatic narrative and documentary have diverged: such extra-diegetic textual commentary is regularly used in non-fiction work on television, never in film.

Figure 5.14.
Abel Gance: *Napoleon*, 1926.
An animated schematic representation of military forces moving across a terrain is introduced into the narrative. in a diegetic context: the viewer sees Napoleon and the other officers look down at this map.
2h 08m.

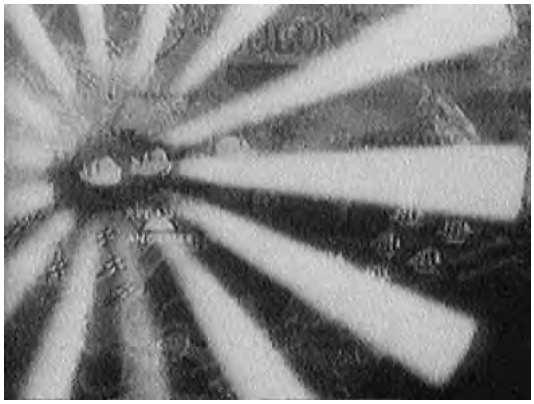


Figure 5.15.
Abel Gance: *Napoleon*, 1926.
On the map are superimposed scenes of battle and schematic representations of conflict.
2h 09m.

In one scene Napoleon and others look at a map, on which an arrow moves signifying the French forces (Figure 5.14). Later the map, some live-action scenes and various schematic representations of conflict are all superimposed (Figure 5.15). As late as 1942, in *Casablanca* (**Curtiz 1942**), such an animated map is superimposed on scenes of travel (Figure 5.16) but significantly this is in the newsreel-style introduction (that is, like the newsreel in Kane it is imitating another genre), not in the main body of the narrative. It appears again in 1981 in *Raiders of the Lost Ark* (**Spielberg 1981** not illustrated) but this also is intended to evoke another, by then historic, style of film-making (**Vaz and Hata 1995 p129**). Animated graphics are entirely acceptable in factual television, but not in a narrative fiction.

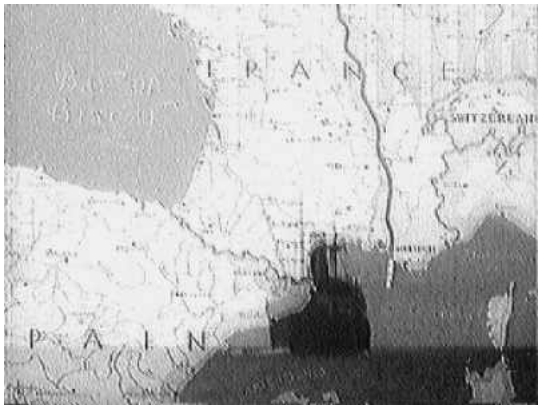


Figure 5.16.
Michael Curtiz: *Casablanca*, 1942.
Superimposition of two modes of information: animated map and live-action scenes of travel.
0h 01m.

Styles of camera movement

Genre also now dictates how cameras may be moved. In Gance's Convention sequence there is an extraordinary swooping tracking shot above a crowd, which is effective in being vertiginous, but seems to have too little relation to natural vision – even to imagined or remembered visual experience (Figure 5.17a-b). It draws attention to the technology.[17]

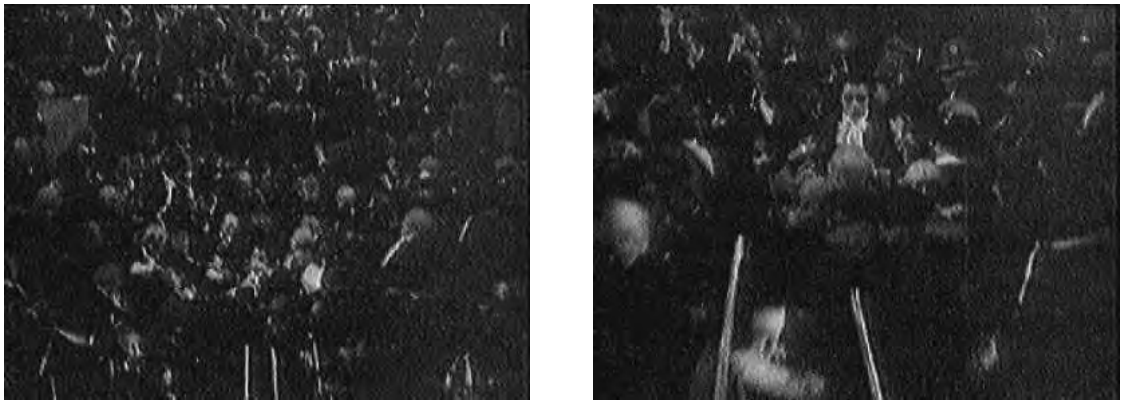


Figure 5.17a-b.
Abel Gance: *Napoleon*, 1926. Parts of the swooping sequence over the crowds in the convention. 1h 46m.

Similar techniques survive in the unmotivated zoom of some music videos, little if ever used in mainstream cinema. A clear differentiation of spatiality on the basis of genre is made by Englander:

> There are two kinds of zoom shot which, in the BBC, we call the 'light entertainment' zoom and the 'drama' zoom, respectively. With the first you must zoom straight in or out regardless of what is is suggested by the subject or scene. But the drama zoom is governed by the tensions within the scene and what is going on in it.
>
> **Englander and Petzold 1976** p112

The techniques used in Napoleon are opposed to those of the subsequent mainstream. Though Harrington claims that the classical film is based on *showing*, it is designed not to be perceived as telling, in the end not even as showing, but aspires to resemble unmediated seeing.

[17] The extensive use of swooping and flying camera in Imax films is gratuitous in the sense that its main purpose is to incite a visceral reaction (and to show off the technology) rather than to further any narrative, but is almost always given some diegetic motive based on the movement of a person or object within the scene.

Overt spatiality: the films of Greenaway

I have shown how the classical fiction film adopts and rejects particular ways of using space according to its objectives. It is not surprising then, that when the objectives are different, so is the space. In the films of Peter Greenaway it is possible to find spatial practices which have been rejected by the mainstream, but which place his films close in some respects to the spatiality of the now rejected styles just described and of factual television. I summarise here the salient spatial characteristics:[18]

There is extensive use of horizontal symmetry, together with the positioning of characters and scenes parallel to the picture plane – both practices which are anathema to the classical film (Figures 5.18 and 5.19).

Figure 5.18.
Peter Greenaway: *The Draughtsman's Contract*, 1982.
Close-up with symmetry.
0h 02m.

Figure 5.19.
Peter Greenaway: *The Draughtsman's Contract*, 1982.
Long shot with symmetry and parallel to the main axis of the scene.
0h 23m.

The films draw attention to the process of representation. For example every effort is made to identify the representation depicted *in* a film such as *The Draughtsman's Contract* (**1982**) with the representation *of* film (Figures 5.20a-b and 5.21).

Cameras (and actors) move along straight lines parallel or orthogonal to the scenery, quite differently from the motivated free movement of the 'eye' in the classical fiction film. This artificial spatiality is reminiscent of a scrolling arcade game.

Figure 5.20a-b. Peter Greenaway: *The Draughtsman's Contract*, 1982. The face-on view of the draughtsman's frames ensures that the viewer becomes conscious of the framing of the photographic shots. 0h 15m and 0h 24m.

[18] Since they are partly formal experiments, his films each explore different kinds of spatiality, but there are many tendencies common to them all.

Figure 5.21.
Peter Greenaway: *The Draughtsman's Contract*, 1982.
The face-on view of the draughtsman's paper identifies the paper with the screen and draws attention to the medium.
0h 06m.

The possibilities are exploited of manipulating and combining images using digital media, notably in *Prospero's Books* (**Greenaway 1991**). The results are *not* designed to be mistaken for straightforward seeing (unlike, say, the use of computer graphics in a film such as Titanic (**Cameron 1997**) which is entirely within traditional Hollywood filmic practices). There is a syncretic approach to modes of representation so that photography, drawing, animation and text are brought together, another feature common in newsreel and factual television but deprecated by the classical fiction film. Frames within frames draw attention to the edges of the image (Figure 5.22).

Figure 5.22.
Peter Greenaway: *Prospero's Books*, 1991.
Frames within frames (and multimodal information).
0h 01m.

Why does Greenaway use a spatiality which is at odds with that of the classical fiction film? Simply because his objectives are different. In his view 'cinema is too rich and capable a medium to be merely left to story tellers' (**Katz 1994 p553**). For him, the exciting areas of film-making are *not* primarily narrative (**Melia and Woods 1998 p130).** While Greenaway's films are engaging, they are also visibly formal exercises. He is *not* interested in hiding the mediated nature of the image. Describing the space of Zed and Two Noughts (**Greenaway, 1985**) Pascoe (**1997 p12**) suggests, 'It is a space unlike any other into which [...] nothing has entered by accident'. This might be a description of any of Greenaway's film spaces, in which nothing is filmed simply 'the way it is' (that is, using the transparent, because conventionalised, methods of the classical fiction film) but draws attention to its own formal qualities.

Space is not just incidental to Greenaway's films but is a subject of interest in its own right. Whereas Ridley Scott (another art-school trained English director of the same generation) uses his strong sense of pictorial space as a means to conventional ends, Greenaway exploits his to engage in formal experiments. He explores ways of objectifying the screen image – making the viewer media-aware – in a way which will always be of only marginal interest to most cinema-goers, precisely because it prevents the psychological immersion which is the essence of mainstream film-making.

Turning from mainstream practice to a film like *Prospero's Books* has shown how hybridity in modes of representation – the mixing of text and picture, of drawing and photography – tends to lead to a medium which is spatially hybrid even within the frame. The configurational possibilities that layering and spatial juxtaposition allow are exploited by Greenaway to construct new relationships not possible when spatial articulation is confined to configuring the model and manipulating the view of it. He gives himself the freedom to use space in the plane as arbitrarily as conventional film uses space in time, but at the cost of losing the simple directness of the unitary image. While time-wise deconstruction and synthesis of space has come to be regarded as natural, to the point where it is unnoticed by the majority of film-viewers, planar synthesis, like that of a cubist painting, will never pass the test of seeming direct and unmediated. To Greenaway this is a virtue not a cost. It nevertheless defines a clear separation between naturalist spatial practice designed to pass unnoticed, and formalist spatial practice designed to be seen.

It is obvious that any identification of one particular spatiality with the *technology* of film must be misleading. Though one set of spatial practices is dominant in film, this is because one genre, the classical fiction film, is also dominant.

7 Film and vision

I have repeatedly touched on the relationship of film to natural vision. I now confront directly the relationship between film and natural vision in terms of both *correspondence* and *function*. In particular I underline again the pragmatic approach which film-makers take and which should inform any attempt to understand how the space of film corresponds to the space of the world.

In the previous chapter, I identified difficulties relating to PI-realism. The first involved possible *incompetence*, but this is not an issue in film, because films are made within a culture where the construction of PI-realist images is well understood, and because film's basis in photography means that many aspects of mapping optical data to the planar image are taken care of by technology. Other difficulties were *implementational*, and I have already indicated that film has similar problems in this respect such as providing the optimal viewing position for the image. However, as with pictures it seems that it is the internal relations within the picture which are important, not the illusion that the picture is a window on a real space.

The other difficulties of realism in pictures were those of *failure of correspondence* and *functional mismatch*. In discussing the question of correspondence in pictures, I argued that imitation of the optical 'sheaf of rays' was only one interpretation of the concept of visual realism, and suggested that most pictures aim instead to capture subjective aspects of vision not accessible in the optical instant. I suggested that many pictures correspond more closely to what it is like to *see* than to the snapshot. The question of *correspondence*, then, is a matter of asking what pictures, or film, correspond to.

Secondly, under the heading of *functional mismatch* I suggested that, even supposing some definitive correspondence to this subjective aspect of vision could be established, there were many reasons why the picture-maker might choose to avoid it if the objectives of the artefact were not well served by such kinds of visual realism, especially if the adherence to some strict form of realism led to a loss of expressiveness.

These two questions are now pursued in relation to film. This is an opportunity to synthesise them into a clear picture of the rationale for the spatiality of film before moving on to other screen-based media.

Correspondence: is film like vision?

There are some techniques which might seem to hold out the promise of being more like vision than others. For example, cutting could be minimised or even eliminated; the camera could move in imitation of the movements of the head and eye, and shots could have maximum focal depth so that the whole frame would be in focus. Such techniques have been advanced as more realistic (meaning more like natural vision) in particular by Bazin (**1967**), an advocate of Welles' innovations in this direction. Part of the difficulty with this argument is that it presupposes what constitutes vision. I noted the dilemma between using uniform and differential focus in pictures, since either is justifiable as being 'like' vision, and the same difficulties arise (and not just with focus) in relation to film.

Film is not like vision in some basic ways. Three examples may serve as illustration. They are offered only tentatively, since they require closer investigation if any definitive conclusions are to be drawn, but they suggest the nature of the problem. They are all concerned with movement, that aspect of film which seemed as if it would be wholly beneficial to visual realism, but which turns out to bring difficulties of its own.

Egocentric awareness

Awareness of one's own movement and unity of position prevents in natural vision anomalies which arise in viewing film. If a camera is positioned under a bridge over a road and views an approaching car, the resulting scene will be as shown to the right of Figure 5.23a. If the camera tracks the car as it passes below and recedes into the distance, the image will be as in Figure 5.23b, in which the whole scene is upside down. This is not however the experience in natural vision. In natural vision, awareness of the body's (especially the head's) orientation prevents the receding car from being seen as inverted. The absence of such feedback, together perhaps with the fact that the image changes under the film-maker's control rather than that of the film-viewer, makes the optical 'truth' of the camera seem incongruous and false.

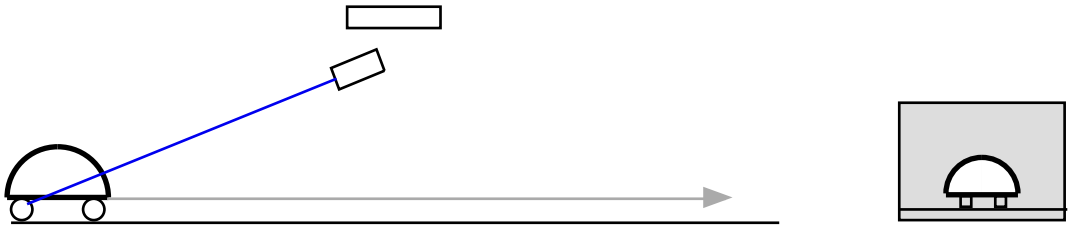


Figure 5.23a. A camera faces a car which approaches it along a road. The resulting image is shown at right.

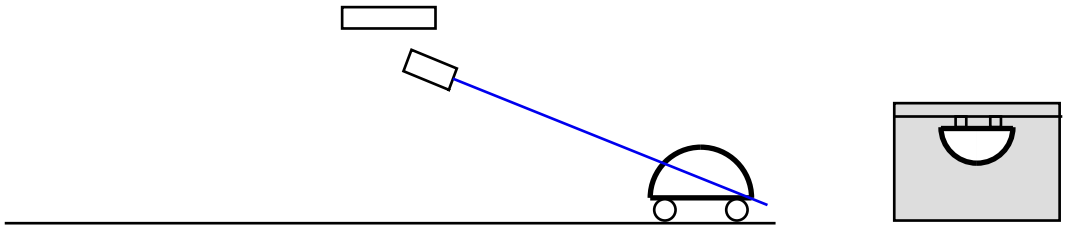


Figure 5.23b. The camera follows the car as it passes below and recedes. The resulting image is shown at right.

Camera movement

The second example resembles the problem of deciding on the relative truth of differential and uniform focus. That difficulty arose because vision takes place in the visual system, not in the retina, and the mind suppresses any awareness of saccadic motion, convergence or focussing. Moreover, it seems that the visual system also partly suppresses awareness of grosser movements. When a hand-held camera is carried though a scene to capture the view as of a person walking, the film-viewer is painfully aware of the wobbling of the resulting image. This is presumably because in natural vision the wobbling of the head is compensated for in a way which does not operate in the mediated form of film. This seems not to be a problem of the inadequacy of a cumbersome technology: on the contrary if a tiny camera could be mounted on the head, or even in the eye, of the camera-operator, the problem would be no less, and probably worse. What happens in practice in film-making when the fluidity and 'situatedness' of the hand-held camera are required without distracting side-effects, is that Steadicam or some equivalent technology is employed to eliminate as much camera-shake as possible. What the film then presents is not equivalent to retinal vision but to the processed mental products of that vision *after* the accidentals of changing position have been eliminated: a clear case of VE-realism at the expense of PI-realism.

Panning and attention

If camera movement were used in place of the conventional cutting between discrete shots, the viewer would know for certain how one side of a room related to another because the camera would have swept across the whole of the intervening space. I discussed earlier the practical argument for omitting such 'in-between' information in the interests of narrative economy. If a camera pans across a scene from one significant detail to another, such as from actor to object or actor to actor, it must take in all that lies in between, and this is contrary to the principle of optimal views, since it fills the screen for several seconds with a great deal of information which is of almost no value. But additionally there is some doubt as to how closely such panning matches vision.

In *natural vision* observers generally close their eyes whenever they make gross changes to their direction of view.[19] One might say that they exercise a jump cut in the visual stream. For the film-viewer to see the sweep of this in-between material is therefore not only functionally unnecessary but also in a limited sense unrealistic. But, since in natural vision the brain eliminates awareness of this blanking during rapid movement, making it seem as though viewing were uninterrupted, it could equally be argued (and often is) that the cinematic cut is *unlike* natural vision. This is another case where there are valid rival claims for what constitutes realism.

Of *Kane* Bazin says, 'Thanks to the depth of field, whole scenes are covered in one take, the camera remaining motionless. Dramatic effects for which we had normally relied on montage were created out of the movements of the actors within a fixed framework' (**1967 p33**). Whereas Bazin habitually claims realism as the rationale for desirable filmic practice, this is an argument for maximal expressiveness – an argument from function, not correspondence. When he says (***op cit* p32**) 'we are witnessing the almost complete disappearance of optical effects such as superimpositions, and even, especially in the United States, of the close-up, the too violent impact of which would make the audience aware of the cutting'[20] his 'realism' has a clear definition. It is not necessarily matching natural vision. The issue is the transparency or otherwise of medium and technique – the need to make the viewer unaware of the medium. This is not now an issue to which any 'objective' realism is necessary: it is defined in its own terms of transparency or otherwise.

Citizen Kane has become something of a battleground in relation to realism. Bazin was the first to suggest that its general avoidance of close-ups, its use of deep focus in which all parts of the frame are clearly resolved, and the long uninterrupted takes which these innovations facilitated, amounted to a style of film narration fundamentally different from the older montage style in which space was fragmented, and specifically suggested that it was more *realistic* (**Bazin 1967 p23-40**). Bazin's arguments are complex, confused even, and in particular he does not differentiate properly between what I have called questions of correspondence – the degree of match to some visual aspect of reality – and functionality – the merits and demerits of particular techniques in terms of expressiveness.

Bazin was writing at a time when, he believed, one was 'constantly being told' that montage was the 'essence of cinema': this in part explains his overstating of the case for techniques which by contrast show a 'straightforward photographic respect for the unity of space' (**Bazin 1967 p46**). As with the appeals to realism in the computer

[19] Evinger *et al* (**1994**) found activation of the orbicularis oculi (the lid-closing muscle) in 97% of saccadic gaze shifts larger than 33º. They eliminated the possibility that these were reflex blinks caused by air-currents over the eye or eye-lashes while the head is turning. The probability of a blink occurring increases with the size of the gaze shift (***ibid* p337**). It is not certain that such blinks serve the purpose of temporarily blinding the eye, since saccadic suppression is also known to occur – the eye becoming partially blind (even when open) during saccades (**Hochberg 1987 p607**): it may be that these 'gaze-evoked blinks' (**Evinger *et al* 1994 p337**) protect the eye during the movement, or simply lubricate it at a time which is convenient in a way that it would not be during fixation (***ibid* p342**). Nevertheless, the phenomenon seems to offer good supporting evidence of 'blinding' between fixations, using one or both of saccadic suppression and blinking, associated particularly with the kind of large-scale gaze shifts for which film-makers have traditionally used editing cuts.

[20] Mainstream film-makers agree that editing should be unnoticed, for example, 'you must never be conscious of going to a close-up or of going to a long shot.' (Henry King interviewed **Brownlow 1968 p109**). Renoir (**1974 p57**) came to consider that 'the best editing is the kind that is not noticed' but like most film-makers he does not make clear why. *Some* kind of realism may be intended, but quite what kind is not explicit.

graphics literature cited in the last chapter, Bazin adduces several different notions of realism. Often he puts the case in moral terms: '*respect* for the continuity of dramatic space' (***op cit* p34**), a technique which 'does not *deform*' reality (***op cit* p27**), 'Murnau has no need to *cheat*' (***ibid***), 'the *tricks* of montage' (***ibid***), and even film's '*vocation* in the service of realism.' (***op cit* p38**, *emphasis added in all cases*).

What is this realism? Of Murnau's films Bazin says that he is interested 'in the reality of dramatic space.' He suggests that 'the composition of [Murnau's] image is in no sense pictorial' (**Bazin 1967 p27**). It seems that the goal is to present the viewer whenever possible with an automorphic mapping of the scene – to extend the full, though unclear, benefits of photography beyond the frame to the shot and to make that shot equivalent to what in a montage-based film would be the edited sequence.

Murnau's work is praised also because of 'the uncompromising realism of a film whose settings are completely natural' (***ibid***). Here it seems that there is a double realism: the scenes are realist in being like real places and the automorphic mapping of photography is transparent to that realism. This is a realism akin to social realism. In Stroheim's films, 'reality lays itself bare like a suspect confessing under the relentless examination of the commissioner of police' (***ibid***); here realism gives access to (unspecified) deeper truths.

Generally Bazin considers realism as being an unproblematic relation to the scene: we know what scenes look like and film should look the same. He prejudges that the thing to which film should correspond is optical vision. However, he also advances arguments which include consideration of *seeing*. An argument for the continuous deep focus shot is that 'it brings the spectator into a relation with the image closer to that which he enjoys with reality. Therefore it is correct to say that, independently of the contents of the image, its structure is more realistic' (**op cit p35**). This is an argument not about the simple matching of the representation to the scene, but of making a representation which causes in the viewer a response similar to that on looking at the scene – the distinction between PI-realism and VE-realism. Bazin goes on to argue that such shots produce 'both a more active mental attitude on the part of the spectator and a more positive contribution on his part to the action in progress. While analytical montage calls for him to follow his guide, to let his attention follow along smoothly with that of the director who will choose what he should see, here he is called upon to exercise at least a minimum of personal choice. It is from his attention and his will that the meaning of the image in part derives' (***op cit* p35-6**). While it should be noted that exactly the same can be said in favour of montage,[21] this seems at least a far more complete view of the problem, and one which takes the argument from matching scenes to making experiences for the viewer. Each technique in its own way offers a set of affordances on the basis of which the viewer constructs meaning. This is what Furnham (**1999**, *abstract*) describes as producing 'an effect on the viewer where the viewer engages in working out what is happening within an array of possible meanings contained within the audio-visual pattern.'

[21] 'By combining these monstrous incongruities we [film-viewers] newly collect the disintegrated event into one whole, but in *our* aspect.' (**Eisenstein 1977** (**1949**) **p34**)

A shot which has proved controversial in relation to Bazin's argument for realism is that from Kane depicting Susan Alexander's attempted suicide. It is now known that great effort was put into contriving an extreme depth of field by matting two shots together, of the drugs on the side table and the figures in the doorway in the background (**Allen 1995 p96**). Similarly Carringer reports (**1996 p94**) that Welles and Toland overcame the difficulties of achieving very long tracking shots by using the optical printer to achieve effects not possible with the unaided camera. It is known that nearly fifty percent of the film was postprocessed using optical printing (**Callow 1995 p522**).[22] This is illusion in the service of 'naturalism.' Carringer remarks that 'Bazin's point is valid, but his underlying premise was wrong: the shot reveals Welles not as a photographic realist but as a master illusionist.' However Carringer here is confusing the means and the result, or difficulties of implementation with failures of correspondence: Welles might simply be using subterfuge to get round the limits of the technology and make what the viewer sees more like natural vision.

When Welles says 'The danger in the cinema is that you see everything, because it's a camera. So what you have to do is to manage to evoke, to incant, to raise up things which are not really there...And the interior conception of the author, above all, must have a single shape,'[23] it is clear that the visual realism of photography is for him a problem as much as a virtue. Welles subverts the concreteness of photography to fulfil his objectives.

Commenting on the work of the Japanese printmaker Sharaku, Eisenstein (**1977 (1949) p32**) says of the nonrealistic proportion of the parts of a portrait face: 'He set up the essence of the psychic expression as the norm for the proportions of the single features.' **Eisenstein 1977 (1949) p33.** In other words Sharaku makes the parts of the face the size they need to be *in order to achieve a certain effect*, pragmatically, rather than the size they actually are. This seems to belong to a functional argument which prioritises expressiveness at the expense of adherence to some external referent. He likens Sharaku's process to that of the filmmaker:

> Is not this exactly what we of the cinema do temporally, just as Sharaku in simultaneity, when we cause a monstrous disproportion of the parts of a normally flowing event, and suddenly dismember the event into 'close-up of clutching hands,' 'medium shots of the struggle,' and 'extreme close-up of bulging eyes,' in making a montage disintegration of the events in various planes? In making an eye twice as large as a man's full figure?! By combining these monstrous incongruities we newly collect the disintegrated event into one whole, but in *our* aspect. According to the treatment of our relation to the event.
>
> **Eisenstein 1977 (1949) p34**

In some ways this is also an argument from correspondence, though not a PI-realist correspondence to an external referent. Just as I have repeatedly noted the possibility of using media to imitate something like the experience of seeing rather than replicating what is 'out there', Eisenstein is arguing that fragmentation allows the

[22] Optical printing uses a combined projector and camera to enable previously processed film to be rephotographed. In the process many liberties can be taken with the original shots: for example two or more shots can be combined, including mixing live action and models, and the film can be rephotographed from a variable distance to imitate zooming (**Katz 1994 p1039**). Carringer (**1996 p99**) reports Dunn (who developed the techniques of optical printing) as saying that in some reels of *Citizen Kane* the percentage of optically printed work was as high as eighty per cent.

[23] A 'recent interview' quoted by McBride (**1972 p39**).

viewer to undertake the reintegration, avoiding the all-inclusiveness of photography, making it a more suggestive and less depictive form. This handing over of the task of integration to the viewer can be considered *naturalistic* in that it may better resemble experience of the world: it makes the film-viewer feel like a natural observer in a real situation which must be assessed, sized-up, made sense of, and less like the passive recipient of an authored presentation.

Even the more prosaic Reisz sees editing as a synthetic process in which fragments of reality are presented to the viewer for *recombination* (**Reisz and Millar 1982 p22**). In one way this recalls the spatial practices of Cézanne, the 'accumulated fusing of little tilted facets' (**Hughes 1980 p27**) and of early Cubism: 'They wanted to compress this inspection, which takes time, into one moment – one synthesised view. They aimed to render that sense of multiplicity, which had been the subtext of Cézanne's late work, as the governing element of reality' (***op cit* p20**). Eisenstein similarly wants to combine views of things seen at different times and scales. Eisenstein – and even Welles to a certain extent – sees the surface appearance of reality as problematic, partly because it is too loaded with specifics and therefore in some ways a poor vehicle for expression. The visible world requires organisation by a stronger principle: 'The dry quadrilateral [ie. the frame of the shot], plunging into the hazards of nature's diffuseness' (**Eisenstein 1949/1977 p40**). The task is 'hewing out a piece of actuality with the axe of the lens' (***op cit* p41**).

What no one could have predicted was that, whereas the forms of Cubism would continue to be seen as artistic interpretation – a strong *interference* with the appearance of the world – similar techniques used in film would become accepted as more or less unmediated representations of looking!

The question of filmic realism – to what does the filmic image correspond? – is I believe *undecidable*. As with pictures, for film there are many alternatives as to which 'reality' is imitated. This problem of correspondence is prior to any consideration of the functions that use or avoidance of realism may offer – how the objectives of a given artefact might be served by deliberately selecting from alternative realisms or by partly avoiding realism altogether. Once a style of depiction has been accepted and absorbed – that is, has become transparent to the content – there is a reluctance to give up what has turned out to be a highly expressive technique. Thus even those films which aspire to seem highly naturalistic are really informed by a deeply pragmatical approach.

In the history of cinema, innovations are tried and if they 'work', they are retained. There is no all-informing theoretical rationale, but a series of tendencies and

preferences – for the classical film above all that of apparent naturalness. As a general rule, nothing should remind the viewer of the characteristics of the medium, such as the flatness of the projected image, the fact that it is captured and projected using mechanical devices, that apparent parts of a unitary space may have been filmed at different times in different locations. It must *seem* unmediated.

We can say that film imitates vision, but in the same ways that pictures do; namely, it aspires to VE-realism – the evocation of visual experience, real or imagined. Even so, if adherence to any kind of realism would confound the objective of constructing a dramatic narrative experience, it is rejected. I have shown that classical films have the form they do as a result of the interaction of the many influences on the variables of the shot, shot selection, shooting style and editing, reconciled above all in a pragmatic way with a view to affording certain effects in the viewer.

The differences between unmediated autonomous natural vision and authorial intention expressed through film-making are fundamental to the spatiality of film. In this mature medium, selected aspects of visual realism are espoused when they serve a purpose, and that purpose is the elimination where possible of any impediments to the sense of 'just seeing'. Such film achieves its objectives by *seeming* realistic in many ways, but it ends up being a strange hybrid of a realism which is optical (based on those aspects of vision easily captured by photography whose automorphism enables it to exploit primal fears, such as vertigo and anxiety at approaching dangers) and high-level realism (based on the psychological experience of vision), even though each of these considered separately would seem to require a different artefact.

The way in which this hybrid has evolved in one hundred years of cinema is a mixture of the almost inevitable (such as the preference for lenses which produce a more or less undistorted image, that is, in which the objects are where they would be if the screen were in fact a window) and other non-inevitable choices which may simply have arisen because an experiment by some particular film-maker seemed to work well in context and became adopted into mainstream practice.

In the end it is not necessary to characterise film as being 'like' anything at all in any pure way. Film may be fundamentally based on the evocation of certain aspects of natural vision, but almost anything else that can be said about film is capable of being contradicted or at least qualified. Film aims to conceal its mediation, yet sometimes visual conceits are used which draw attention to artifice and authorship. Film aims to be like natural vision, yet it often resembles visual imagination or recall. Film should seem real, but the fact that it is a representation is part of its attraction. Film relies on the verisimilitude of photography, but the unselectivity of photography is its enemy. The elements of the classical film have multiple rationales: there is no simple explanation.

The fact that many innovations in film space were makeshift inventions which have subsequently been adopted into mainstream practice – and that the success of so

many of film's spatial techniques could not possibly have been predicted – gives pause for thought in relation to new media. It suggests that substantial open-ended experiment is called for and that the spatialities of new media may be quite other than we currently imagine them.

I have shown that mainstream film is a spatial hybrid, highly automorphic within the frame, but between shots subject to wide variation. Sometimes the inter-shot relationship is approximately automorphic, such as when two consecutive shots look in slightly different directions from the same point. Generally however, between shots belonging to a single scene, it is not automorphic in that simple way, using many different views to give just sufficient illusion of a coherent space to serve the purposes of the narrative drama. I have shown that the time-wise juxtaposition of particular views is strongly motivated by the demands of storytelling and is inclined to be configurational in character: it puts shots in relation to one another under authorial control in order to create meaning, not to capture some pre-pictorial world. In cutting between scenes it is wholly configurational, imposing a structure which yields new meanings not inherent within the elements themselves.

6 Screen Space II: Factual Television

1 The spatiality of television

Several authors have dealt, at least in passing, with the spatiality of mainstream film. The same is not true of television, whose spatial practices are almost entirely unanalysed. Books such as those of Merritt (**1987, 1993**) and Crook (**1986**) celebrate the inventiveness of the graphic designer for television in creating title sequences and channel idents but have nothing to say about the overall spatial qualities of programme content.[1] Gessner (**1968 p272**) dismisses TV space as 'visually disorganised.' However I will show that the spatiality of television *is* organised when considered in relation to its objectives. The focus is factual programming, using examples of news, documentary and adult education, in order to identify the distinctive issues.

There is a far wider range of genres within television than in mainstream cinema, for example sports, light entertainment, education, arts, drama, soap opera, music-tv, situation comedy, advertisements, stings and idents, news and current affairs, though some of these named genres interpenetrate. Considered together they present a gamut of spatial practices extending from those broadly similar to the classical fiction film – for example in costume drama or soap opera – to those which are very different. It is in relation to the latter that I show how many of the spatial features of the newsreel which were summarised in Table 5.09 (in the previous chapter) have survived and indeed flourish in factual television, so demonstrating that spatial practices are selected for their fit with the objectives of the artefact, not on any universal basis. It is their very opposition to the spatial practices of the classical fiction film which makes them worth special consideration. However, as might be expected, even within the genre of factual programming there is some diversity of practice, so as always what is documented is a series of tendencies rather than a uniform behaviour.

In the absence of any literature or record of past practice, to what extent can the spatiality of current factual television be described without seeming to discern broad principles on the basis of short-lived, perhaps merely fashionable, practice? I have shown how a century of film-making practice has refined the spatiality of the fiction film such that, though innovation and change will continue indefinitely, the fiction film can be claimed as a mature medium which has achieved a high degree of expressiveness: its spatiality is so well attuned to its objectives that it is largely resistant to the sudden variations of fashion. While television has itself been developing for half a century, it is difficult to discern practices which are acquiring any level of permanence. In the analysis that follows, it seems evident that to a far greater extent than in the fiction film, change itself is a norm. The ephemeral nature of much television broadcasting and its need to be seen to be constantly new have a significant effect on spatial practice. I noted how spatial innovation was in itself

[1] A chapter in Crook's book on the functions of 'identification, explanation, promotion [and] props' proves to contain nothing at all about explanation and to be entirely given over to identification and promotion.

problematic for the fiction film in that it risked drawing attention to the means of representation, but this is not a difficulty for factual television, especially in peak-time programming such as popular documentary and news broadcasts, whose creators by contrast want to be noticed as innovators. It is important therefore to suspect that any particular cluster of spatial practices associated with a particular genre may change in time far more radically than those of mainstream film. Provided this characteristic is recognised, it is still possible to show general spatial tendencies which make the genres distinctive.

The three factual television examples discussed – an Open University adult educational programme; television news broadcasts on BBC News 24, Sky News, CNN, Bloomberg TV and Channel Four; the popularising science programme Superhuman (BBC Television) – are not uniform in their approach. Even within one genre they represent a gamut of approaches; for example Superhuman borrows in interesting ways from the practices of advertising, while the more sober Open University programme is clearly designed for repeated showing over many years. Nevertheless, it will prove possible to find many spatial characteristics which unite this genre in opposition to the practices of mainstream film.

I have shown that the classical fiction film, when taken as a whole rather than in terms of individual shots, cannot be conceived as 'realistic' in any sense of mapping the world. Nevertheless it is *naturalistic* in the sense that the medium is intended to be 'seen through': it aspires to seem largely unmediated. The distinctive spatial qualities of factual programming for television can be roughly characterised by their opposition to these concerns and by their inheritance of spatial practices originally invented in film but subsequently expunged. However, even within this genre, it will be apparent that there is significant variation of spatial practice – in particular in the way in which two complementary spatialities are exploited. These are the space of the *model*, in which configuration occurs pre-pictorially (or at least appears to do so) and the space of the *picture*, in which viewing and picturing are the main means of configuration.

I indicated in the last chapter how closely in film the manipulation of the model ***M***, the view ***V*** and the picture ***P*** are interrelated, and do not want here to suggest that they are separate: the difference between model-spatiality and picture-spatiality is a difference of emphasis only, so that any given programme or programme segment will tend to achieve its objectives through greater use of one or the other. For brevity's sake I summarise the spatial characteristics of each of three different kinds of factual broadcast, emphasising what they have in common (and which differentiates them from fiction film-making) but also identifying whether they achieve their objectives principally by manipulating ***M*** or some mix of ***V*** and ***P***.

The manipulation of pictorial space in both television and multimedia includes the use of basic configuration – juxtaposition, grouping and alignment – of multimodal segments of which pictures are often only a part. Having seen how the space of film is

only automorphic within shots, but configurational between them over time, we are now dealing with images which are often configurational in space as well.

2 Non-fiction television space

For television it is again clear that while technology may set the outer limits of what can be done, it is *genre* which determines which spatial practices flourish and which decay. Ellis rather surprisingly suggests (**1992 p159**) that 'there is no real difference in narrational form between news and soap opera'. Whether or not this is true of narrative style, it can certainly not be said of the spatial practices of the different genres. While the soap opera broadly imitates the spatial characteristics of the classical fiction film, factual television could hardly be more different.

Example one: Open University adult educational programme

In an Open University programme made some time during the early 1980s, the following spatial characteristics are noticeable:

Presenters appear against a plain studio background, facing the camera – and therefore the viewer – unlike the normal practice of film (Figure 6.06).

Figure 6.06.
Open University: *Electrostatic Charges*, early 1980s.
Simple, direct-to-camera presentation against a plain background.
0h 00m.

Again Ellis treats all television as equally different from film: '*film* events do not betray a knowledge that they are being watched' while *television* is 'forever buttonholing, addressing its viewers as though holding a conversation with them.' (**Ellis 1992 p132**, *emphasis added*). He overstates the case, since, as indicated earlier, many genres of television aspire like the classical fiction film to conceal their mediation, but his remark captures the character of much factual broadcasting.

Figure 6.07.
Open University: *Electrostatic Charges*, early 1980s.
A crude painted equivalent of the cloud scene has schematic graphics overlaid on it.
0h 00m.

Disparate modes of information are overtly used together: film, diagrams, still

photographs, animation and text.[2] I have noted the unacceptability of such mixed modes of representation in the fiction film.

Film is speeded up to exaggerate the processes depicted and make economical use of time. By contrast fast motion is rarely used in mainstream film, except covertly as a subterfuge to make an action seem faster than when it was filmed. Its overt use in this broadcast inevitably draws attention to the medium. Historically it has been a comic device, and comedy often seems to involve a kind of objectification.

Figure 6.08.
Open University: *Electrostatic Charges*, early 1980s.
Speeded live-action film of a cloud accumulating.
0h 00m.

Unrealistic configurational relations are constructed between disparate elements. A presenter is seen in company with a graphic, so that she may comment directly on it (Figure 6.09). On referring to a change in the state of the cloud (visually indicated by new symbols appearing) she turns to look at the change she is describing, making use of the viewer's tendency to observe what others look at noted earlier. This is a borderline case between manipulation of an apparently pre-pictorial model and pictorial configuration. It seems as though the graphic really is there with the presenter, yet the viewer probably does not see it as the wall-sized object it would be if it were really in the space occupied by the presenter.

Figure 6.09.
Open University: *Electrostatic Charges*, early 1980s.
The presenter looks at the graphic material in order to direct the viewer's gaze, even though the graphic is probably not really visible to her.
What size is this graphic – as tall as the presenter, or the size of the screen?
0h 03m.

To show two things at once, the screen is simply split (Figure 6.10a-b). The relation between the activities in the two parts is important, and there are none of the fiction film's reasons for rejecting this method of solving the problem. This configuration is clearly confined to the pictorial component of the space, just the kind of pictorial configuration which has been expunged from mainstream fiction film-making.

[2] By using different media, different realisms are exploited selectively for their own strengths, with pragmatism again to the fore, selecting *ad hoc* those realisms which best support the objectives of each component. There is no role here for the classical fiction film's sense of a single unitary realism.

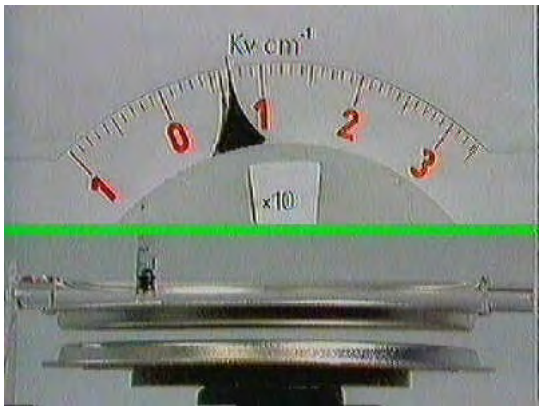



Figure 6.10a-b. Open University: *Electrostatic Charges*, early 1980s. In ***a*** a medium shot of the experimental set-up establishes the spatial relation between the parts. In ***b*** the most relevant parts of the equipment are forced into proximity using split-screen. 0h 06m.

This programme is certainly a narrative, so the spatial differences from the classical fiction film cannot be explained by the absence of narrative continuity: basic narrative spatial practices like those of film are in evidence, such as the use of a long shot to establish the spatial relationship of the parts of the experiment. The explanation is not a different narrative structure as such, but the lack of a requirement that the viewer forget the process of representation.

There is a marked tendency to use basic configurational and pictorial devices, ***V*** and ***P***, to convey meaning, in contrast to the fiction film where configuration is made to *seem* a property solely of the scene, the model ***M.*** It will be seen later that not all television, even within the broad genre of non-fiction, opts so clearly for this approach.

At each moment, the spatial organisation is optimised to convey the necessary information, sometimes wholly filmic, sometimes diagrammatic, sometimes through the spatial juxtaposition or superimposition of multiple components. The programme's spatiality is largely determined by simple communicative requirements which are answered in the simplest and most economical way.

Example two: Television News
(BBC News 24, Sky News, CNN, Bloomberg TV, Channel Four)
Television news provision is an increasingly competitive business. Considerable resources are devoted not only to gathering news but to its presentation. Broadcasters 'show off' what they can do technically and aim to make their programmes visually appealing. Though each of the broadcasters wants to be distinctive, there are many characteristic spatial factors which they have in common:

Spatial complexity Television news is becoming increasingly complex in spatial terms, both in terms of division of the plane and in the use of multiple planes in depth. Live-action, pictures, diagrams and other media are juxtaposed.

Complex textual interventions The simple caption used historically has developed into a complex pattern of textual overlays and insertions. This emphasis on text brings the news broadcast strongly into the ambit of simple spatial configuration while

continuing to also rely on the pictorial and filmic spatial practices described in Chapters 3, 4 and 5. These television genres are moving from a display which transparently presents the unitary image of cinema to displays dominated by configuration of parts.

Multi-layered ambiguous spaces Some of the distinctive spatial character of current news broadcasting is facilitated by digital technology. However, many of the spatial characteristics of factual television were in evidence before digital media made them easy to achieve. In a pre-digital election broadcast (Figures 6.11a-b) cut-out cardboard shapes were applied by hand in front of the camera, but despite using real objects in a real pre-pictorial space, the siting of these objects in a multi-layered and somewhat ambiguous space looks forward to the spatialities which would subsequently be made possible by digital techniques. The ambivalent relationship to information based on an assumption that many viewers require motivating by decoration, including intriguing spatial puzzles, is already in evidence.

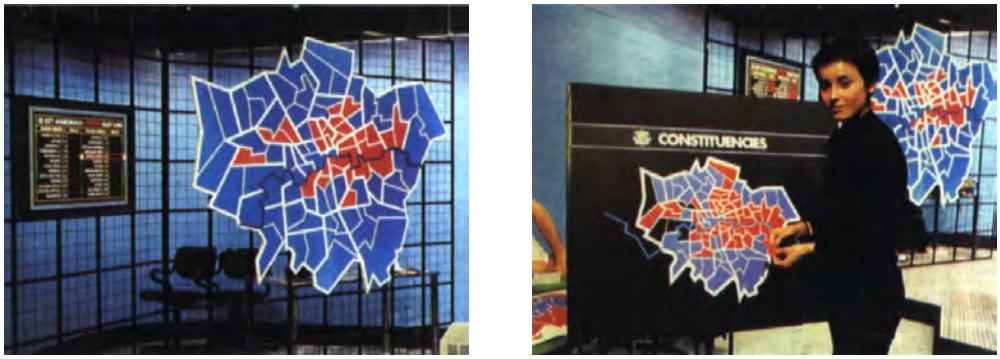


Figure 6.11a-b. Described as 'the old form of studio presentation ... of a London Borough election some years ago [at 1987].' From **Merritt 1987** p86.

Pictorial frames and filters Many news broadcasts carry a logo and other fixed overlays which are attached to the viewing frame rather than to the material being viewed. An example from BBC News 24 (November 2000) has such information in every corner of the screen, augmented by captions (Figure 6.12) and other devices. These objects belong to the picture, not to the view or the model, which are seen as beyond or through them. Historically such graphic objects were common in oriental pictures (eg. Figure 3.19, Chapter 3); now they are a distinctive part of the spatiality of computer games (Figure 3.28, Chapter 3). They are becoming significant in other kinds of interactive pictorial media, reviewed in the next chapter.

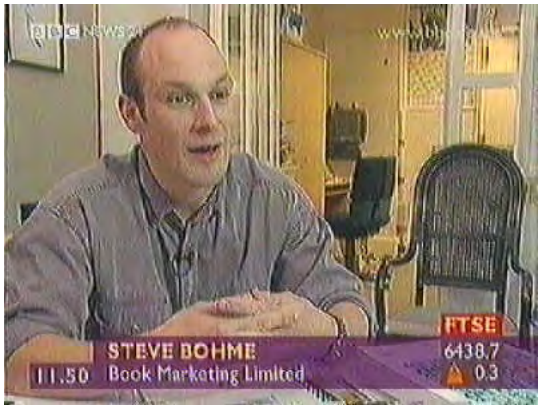


Figure 6.12.
BBC News24, November 2000.
Overlays plus a translucent caption layer.

Ad hoc reconfiguration The segment displaying the presenter may be shrunk at need so that this smaller window becomes a component of a larger display (Figure 6.13). What was a more-or-less cinematic visual world is suddenly reduced to the status of a segment in a non-figurative composition.

Figure 6.13.
BBC News24, November 2000.
The full-screen display is shrunk (with a cut) to a subsidiary role.

The acceptance by the viewer of this sudden scaling is probably built on the familiarity of windowing systems in desktop computer interfaces, but it also harks back to pre-Renaissance configurational approaches to pictorial information, before the notion of the unitary perspectival world became regarded as the norm (and still is in the classical fiction film).

Figure 6.14.
BBC News24, November 2000.
The main multimodal display, including its presenter, is shrunk so that it may be presented by a signer.

Pictures within pictures In addition to pictures within schematic configurations, there may be pictures in pictures, so that the whole of the normal display is shrunk into a smaller area. This inner area has an ambiguous character and relationship to the full screen, at times seeming a miniature within the plane, at others a space receding in depth (Figures 6.14 and 6.15).

Figure 6.15.
BBC News24, November 2000.
The relationship between the signer and the subject takes on an almost pre-Renaissance spatiality.

Multiple views by pictorial and pre-pictorial means Both pre-pictorial (***M***-based) and pictorial (***VP***-based) configuration are used as ways of presenting dialogue between a presenter and others at remote locations. A generally pre-pictorial solution is to display the other party in a monitor on the studio wall. This allows the newsreader to be seen facing the other person, situating the conversation in pre-pictorial space (Figure 6.16). However, while the remote individual is speaking, the whole screen is taken periodically over by a standard-sized close-up of the speaker, so that the pre-pictorial and the pictorial become exchanged. The oddness of seeing a person talking to a representation on a wall passes largely unnoticed because functionally the configuration is expressive – it fulfils its informational purpose.

Figure 6.16.
BBC News24, November 2000.
The presenter converses with a representation on the wall.

A more pictorial solution is used in Sky News (Figure 6.17) where the participants are presented in a diptych (a perhaps unwitting revival of a pictorial structure which might have been thought defunct). The illustration also shows the use of independent captions and tickers for unrelated news stories, a segmented approach which is taken much further by Bloomberg (discussed next).

Figure 6.17.
Sky News, November 2000.
A diptych allows both parties to a conversation to be seen. One is grounded in a known location, Manchester, while the other is in a virtual place, Sky Centre.

Independent information streams Bloomberg's broadcasts (Figure 6.18) come nearest to the segmented space of the computer's graphical user interface: there are at least seven streams of information on screen at any one time and, unlike those of most news broadcasts, they do not belong to a single story but are independent of one another. There is no intention that the user should integrate all the streams of information, though some aspects of the information could probably be monitored in

peripheral vision while concentrating on another.[3] There would be nothing to stop the different segments from being supplied by more than one broadcaster.[4]

Figure 6.18.
Bloomberg, March 2001.
Multiple independent segments in a single display. The would-be immersive becomes objectified when confined to just one segment among many.

A side-effect – which is predictable in light of the differences between unitary cinematic displays and segmented configurations already discussed – is the objectification given to the components. The top left part of the screen is occupied by a standard televisual channel which at times contains advertisements; instead of these engaging and absorbing the viewer in a subjective experience, they are objectified, their swooping and whirling spaces becoming noticeably incongruous in the context of the rest of the screen-space. This is partly a consequence of the positioning of the televisual stream: if it were in the centre it would be easier for it to capture the viewer's attention: all computer games and virtual worlds faced with the issue of accommodating an engaging environment and ancillary information give the centre of the display to the immersive segment.

Presenters embedded in graphics During BBC News24 weather forecasts, the screen is dominated by a series of animated maps, digitally composited to appear as behind the forecaster (Figure 6.19). Since the presence of the foreground graphic layer is maintained, the presenter is effectively embedded between two layers of graphics. As

Figure 6.19.
BBC News24, November 2000.
The presenter sandwiched between graphic layers.

[3] For example the average colour of the two tickers, in which rising prices are green, falling are red, and unchanged are white, tends to indicate the overall state of the markets even when not being attended to.

[4] A recent article by Fischetti (**2000** p33) suggested that the temporal insertion of advertisements into television broadcasts would, when users can digitally omit advertising, be replaced by spatial insertion in the form of banners and overlays. Examples already exist of three-dimensional digital inserts of advertising material into sports broadcasts.

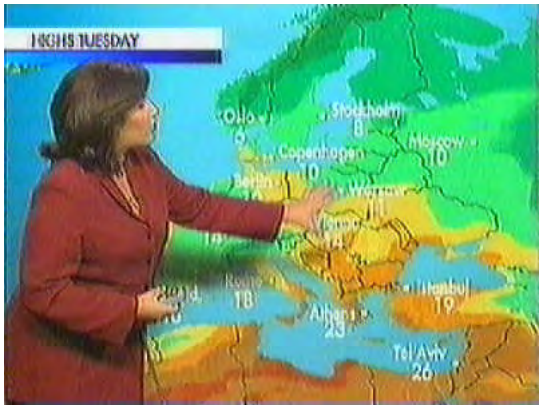


Figure 6.20.
CNN News, November 2000.
The presenter casts a shadow on the virtual weather map.

these layers are positioned in the pictorial space of the screen, presenters seem to be partly dematerialised, an effect augmented by the fact that they move only in the plane and never orthogonally. Unusually in CNN News, the weather forecaster casts a shadow on the maps she describes (Figure 6.20). This evidence of materiality is in contrast to the general tendencies of the genre.

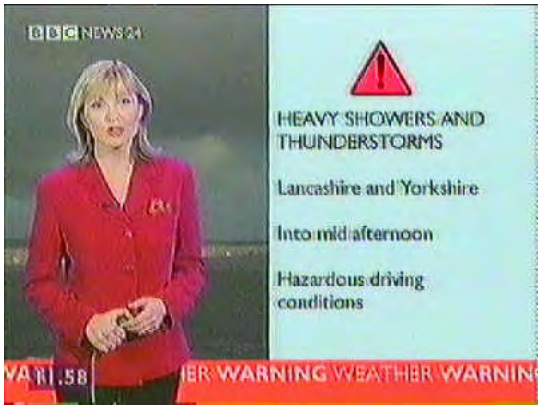


Figure 6.21.
BBC News24, November 2000.
A layered and segmented space to support multimodal communication: the studio presenter, a static textual warning, a scrolling textual warning, live-action scenes and the usual overlays.

Further *ad hoc* changes to the display are called for when weather forecaster, textual information and filmic evidence of the weather must be combined (Figure 6.21). In the BBC News24 example illustrated a further element is the scrolling ticker of weather alerts at the bottom of the screen, making four concurrent streams of information in all.

Figure 6.22.
Sky News, November 2000.
Presenters at Sky Centre, with monitors showing connectedness to the outside world. Whether this apparent pre-pictorial space exists or not is unclear.

Interchangeable virtual spaces and places There are many examples where it is unclear whether a space really exists pre-pictorially or not. Frequently news-readers are seen with various television monitors visible at some distance behind them in the studio showing additional, but generally indecipherable, images (Figure 6.22). It is not

always clear whether these are really situated in some physical wall or in a virtual pane of the screen.

Archaic cinematic techniques In some channels the use of transitions is in the tradition of the cinema newsreel pastiched by Welles. Figure 6.23 shows a horizontal wipe between the image for two stories.

Figure 6.23.
Sky News, November 2000.
One trailer for a news story is replaced by another using horizontal wipes, long expunged from naturalistic film-making.

Also reminiscent of techniques defunct in the fiction film is the use of extended superimpositions, currently favoured by CNN (Figures 6.24 and TV.25). As with so many other characteristic spatial practices, the highlighting of the mediated nature of the image is not the problem for factual television that it is for the fiction film.

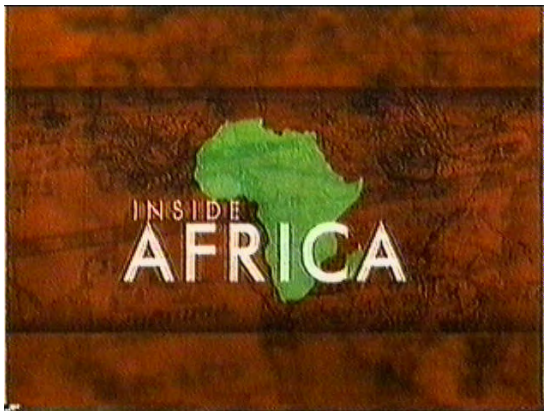


Figure 6.24.
CNN News feature: *Inside Africa*, November 2000.
Superimpositions of lettering, maps and live-action are used.

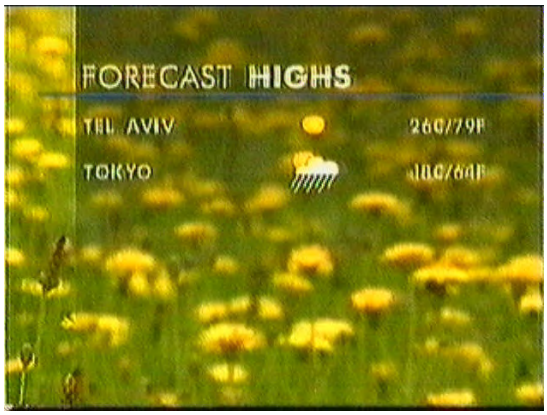


Figure 6.25.
CNN News Weather Forecast, November 2000.
Superimpositions of lettering and symbols on live-action landscape scenes.

As with all the spatial practices discussed in this thesis, those of television have both informational and affective objectives. Together these explain its characteristic form.

In terms of informational expressivity, television news must work within important limitations. There is much less control over the material than is available to the filmmaker. Unlike when making pictures or fiction film, there is little opportunity to configure the model ***M***, the depicted scene, to articulate the meanings intended. It is to a large extent a given, and not designed symbiotically with the view and the depiction as in classical film. Each filmed segment has its own accidental visual characteristics. In addition, the material is by its nature fragmentary, making it difficult to immerse the viewer in an engaging experience. Finally, the material comprises multiple modes of information, including the inherently schematic mode

of text. Together these limitations are bound to push the genre more towards pictorial and configurational manipulation than is the case with the fiction film and its related television genres.

In relation to affective expressivity, there are again a number of distinctive factors. Visual patterning for its own sake – both over time and spatially – is considered important, because the programme must be formally appealing regardless of the interest of its content. Such formal characteristics were noted as an important influence also on the fiction film, but here they are allowed freer rein over spatial configuration. The design of television news broadcasts must reconcile a desire for affect, particularly with a view to suggesting 'liveness', connectedness, and variety, with the need to be seen as a reliable source of unbiased information. The great majority of segments in a factual television broadcast are photographic because of photography's special authority as a means of conveying what is occurring or has occurred, because they impart a sense of immediacy, and because film and photography are relatively quick, effortless and inexpensive means of depiction (compared for example with digital models or hand-drawn animation).[5] Factual television does not aim as the fiction film does to make the user mistake showing for seeing, yet within each of the time-wise segments and divisions of the plane the inherent believability derived from the automorphism, authority and immediacy of photography is extensively exploited. Though the space must perform interestingly, it must not seem to augment or modify the content in ways which would make it unbelievable.

In current news broadcasting practice, while it seems likely that the multiplanar character and the resizing and reconfiguring of panes at need (though by the system not the user) are derived in part from the graphical user interface, there are important differences. In the standard computer interface there is no necessary relation between any two things on the screen, whereas here the separate planes and areas are (with the exceptions noted) contributory strands of a single theme, and what is more are part of an authored narrative. As a result, in basic informational terms the depth-wise ambiguity favoured by most channels, though it seems to serve no practical purpose, is at least not actually detrimental. Basic gestalt principles, assisted by learned convention, ensure that the relationships between the parts are understood.

Clearly, the spatiality of television news, though it is designed to be engaging, has no need to make the viewer forget the nature of the medium as the fiction film does. In fact, that each segment is objectified is beneficial to the viewer-news relationship, just as much as it is disastrous to the psychological immersion of the fiction film. It could be argued that if viewers are aware of the multiple segments then they appreciate the wealth of information with which they are supplied.

[5] The truthfulness of television factual film became a point of legal contention during 2000 in a dispute with specifically spatial characteristics, the opposing positions being that of a journalist, Thomas Deichmann, for the magazine Living Marxism that 'the image was created by "camera angles and editing" and "that there was no barbed wire fence surrounding the camp..."' and that of ITN (Independent Television News) as summarised by a Guardian correspondent, Edward Vuilliamy, that "ITN filmed that which was before our eyes. The prisoners were there, the fence was there." (The Guardian, 21 February 2000).

The overall result is a spatial form which – however one might criticise its often gimmicky and gratuitous visual complexity – is highly attuned to its objectives. In every respect it is possible, on analysis, to see why it has the form it does. The spatial form and the objectives are well matched.

In the next example, a lavishly funded documentary for prime-time television aimed at an international audience, it is also true that the spatial practices adopted closely fit the programme's objectives. While some of the core requirements which informed the Open University programme persist – of presenting information in direct, easily comprehensible form – the balance is tilted even further from informational expressiveness towards affective considerations by the need to maintain interest throughout fifty minutes (and this interest must also be carried over to the next week's viewing figures, since this is a programme in a series). These requirements will be seen to have a powerful influence on the spatialities adopted. In particular, considerable effort is put into making pictorial configuration appear *not* to be configurational but to belong instead to the model. This is done in order to present a broadly educational, factual programme without the characteristic multi-segmented, multi-modal, objectifying interface which is seen to be normal for such material, and which the programme makers presumably feared would be literally a 'turn-off'.

Example three: Superhuman, BBC Television, 2000
For this analysis I have selected one characteristic programme from the series, together with other related artefacts selected (significantly) from television advertising.

Formal complexity as entertainment The programme opens with a short title sequence of brief visually striking shots at a variety of angles and scales and which combine representations with different scales and modes, such as natural photography and x-rays, models and live-action. Their briefness and visual contrast makes an almost abstract time-wise pattern while establishing roughly what the programme is about. The informational requirements of such material – idents, stings and adverts – are slight and allow the designer free rein, often with a substantial budget. The dominant spatial characteristics of such sequences are extreme spatial depth, strong contrasts of scale and spatial ambiguity. Many sequences make use of more or less surreal spatial conceits such as an apparently unmediated photographic scene which is suddenly disturbed like a pool by a stone (Figure 6.26), apparently insubstantial lettering which is shattered by piercing arrows (Figure 6.28) or a globe which opens to reveal a series of internal devices (Figure 6.27). Not surprisingly given that text is the only compulsory element of any title sequence, lettering is frequently used as a primary graphical component, animated, modelled, used as a window on images, morphed into non-textual objects, and in many other forms (Figure 6.29). Stasis and flatness are avoided at all costs. Such traits are also in general characteristic of *Superhuman*.

Figure 6.26.
Diverse Television: *Showreel: 'Lakes' programme title*, 1999.
Lettering responds in a watery way to the fall of a droplet.

Figure 6.27. Pirate: *Showreel: Title for weather forecast*, 1997. Paradoxical spaces within spaces. A globe opens to reveal a series of interlocked, light-emitting mechanisms.

Figure 6.28.
Pirate: *Showreel: Advert for headache remedy*, 1997.
Lettering which at first appears virtual is revealed to be physical when shattered by arrows.

Figure 6.29.
Pirate: *Showreel: Ident for Channel Four Television*, 1997.
Lettering takes on a physical form, though framing and camera movement remain fluid and anti-naturalistic.

Figure 6.30.
BBC Television: *Superhuman* presented by Robert Winston, 2000.
The presenter is composited with a fast-motion shot of fungus growing.
0h 00m.

Scale and depth, pictorial and pre-pictorial In the opening sequence of *Superhuman* which follows the title, a shot appears in which Robert Winston, the programmes' author and presenter, is seen walking in a wood while in the foreground fungus grows (Figure 6.30). The scale of the background and foreground scenes is consistent with their really coexisting in a pre-pictorial space, but any illusion that they do is undermined by the fact that they are in different timescales – the fungus is seen growing. The technique serves the practical purpose of showing both presenter and subject matter, but of course the reasons for its spatial form extend far beyond that requirement. The principal objective is to intrigue, perhaps amuse, and this is served in two ways. The digital compositing of two scenes of such contrasting scales is something of a novelty, and for this kind of programme that is in itself sufficient motivation for its use. But in addition there is an appeal in strange juxtapositions of scale, if those differences of scale appear to arise out of natural vision, which can be traced back to some of the earliest experiments with geometric perspective. Alpers (**1983 p83**) notes the fascination with scale of Dutch seventeenth century painters who 'juxtapose a bull or a looming cow against a tower made tiny by its distance.' She also notes (***op cit* p22**) the enthusiasm for the loss of a sense of 'true' scale as a result of using microscopes and telescopes (which earlier times had regarded as a failing). Such incongruities are repeatedly exploited in *Superhuman*.

Spatial punning Continuities which in the classical film create an illusion of pre-pictorial unity of space, are here used mischievously, suggesting and then undermining spatial coherence. Winston picks a bluebell in a wood (unsurprisingly seen growing at an accelerated pace); when he is next seen holding the bluebell it is obvious that he is in an altogether different location. The presentation is designed again to bring out its own artifice. The viewer is not intended (presumably) to disbelieve the informational content of the programme through analogy with the obviously deceptive means of presentation, so one must conclude as always that the viewer's inference of intention is vital: the viewer, experienced in the spatial practices of television, recognises – and thereby deals with – the divergence between the duplicitous mode of presentation and the apparently honest intentions of the author-presenter in relation to the content.

Configuration made to seem to belong in the model Further games are played with scale and spatial coherence to convey factual information using counterfactual modes of presentation. As Winston speaks, situated in a real-looking if somewhat improbable location resembling a warehouse, he is startled by the appearance of a giant lizard, whose sweeping tail for an instant threatens to strike him (Figure 31a-b). This leads into a more or less realistic use of the same computer-modelled lizard to illustrate the reptile's ability to regrow a severed limb. Needless to say, the conjunction of the lizard and the presenter in the same apparent space serves no straightforward informational purpose. The effects of the persistent use of such techniques in the programme seem to be: (1) to situate the presenter *in* his subject matter, whereas simple cuts would present him as an external commentator;[6] (2) to contrive a

[6] In another programme in the series, Winston is seen looking at and commenting on a medical phenomenon as depicted in a historic oil-painting, but subsequently he himself enters the 'painting' and interacts with the characters within the pictorial space.

continuity between discrete parts of the programme which might otherwise seem fragmented and episodic, creating a synthetic spatial coherence in the absence of a real one; (3) to amuse and intrigue the audience and motivate them to continue watching.


Figure 6.31a-b. BBC Television: *Superhuman* presented by Robert Winston, 2000. The presenter is composited with a computer graphics shot of a lizard regrowing a limb. 0h 27m.

Where the Open University broadcast accumulates fragmentary elements with only just enough spatial coherence to make it clear where a close-up belongs in a larger pre-pictorial space, *Superhuman* frequently contrives to spatially unite things which really belong in different spaces and scales. While using a range of representational media such as live-action, archive film, computer graphics and model shots, it aims to unite these elements in a virtual three-dimensional space whose coherence is perturbed but not interrupted by changes of material.

Text and flat graphics are almost entirely avoided, since they would risk reinforcing the planar character of the display (and because they would have unfortunate connotations of educating and informing rather than entertaining). Thus configuration is made to seem a property of the pre-pictorial model in preference to configuration at the pictorial level but, ironically, this pre-pictorial space does not exist, being engendered largely by pictorial means!

Though the programme makers are probably right to assume that viewers will distinguish between the tricks of representation and the honesty of the content, it is a somewhat risky strategy and would be considered inappropriate in a programme such as news or current affairs in which probity must not seem to be undermined by other motives. There would also of course be prohibitive problems of both time and expense.

Except for the important shared reliance on the realisms of photography, the spatiality of factual television can be broadly defined by opposition to that of the fiction film. Where one is unitary the other is multi-faceted; where one gives transparent access through psychological engagement the other patently presents its content; where one coordinates fragmentary views to construct apparently coherent space the other uses many disjoint glimpses; where one is unimodal the other is syncretic; where one favours the diegetic the other uses extra-diegetic textual (and oral) commentary. However, I have shown that a programme like Superhuman, while it wants to be noticed in terms of the variety and richness of its components, tries in part to shift back towards the unitary image of an apparently preexisting model familiar from the classical film.

Many practices of the factual television genre pragmatically combine automorphic and arbitrary characteristics of scale, perspective and mode of representation in ways which recall pre-Renaissance pictures.[7] Crary (**1990 p2**) suggest that 'most of the historically important functions of the eye are being supplemented by practices in which visual images no longer have any reference to the position of an observer in a "real" optically perceived world.' Though this is a questionable assessment of displays such as fiction film, it is a reasonable summary of the spatial practices of much factual television, and will be seen also to describe well the spaces of interactive multimedia.

Factual television's combination of the figurative-pictorial with schematic and textual configuration, bringing together two different forms of spatiality, will be seen to be also the basis of many interactive multimedia artefacts. Both news television and many forms of interactive multimedia aim to benefit from a combination of two kinds of 'just seeing' – segments which seem to offer unmediated vision, embedded in simple structures which aim to also make the interrelationships of the segments instantly accessible. Nevertheless, as the arguments of this chapter have already made clear, such hybrid configurations cannot, taken as a single display, achieve the illusion of unmediated viewing which is the hallmark of the fiction film. The intervention of configuration within the display is fatal to the ability to forget that the image is mediated. Where instead an immersive and apparently unmediated visual experience is to be offered then, just as with cinema and many genres of television, it seems that only a single, apparently mimetic, image is acceptable.

[7] To treat the Renaissance as a simple watershed is an oversimplification of the history of depiction. Alpers points out there was still, even at the height of the northern Renaissance's concern with mimesis, a willingness to combine modes such as text and depiction, modelling and mapping (**Alpers 1983 pp169, 172 and 197**).

7 Screen Space III: Depiction and the space of Interactive Media

1 Introduction

The objective of this chapter is to consider digital interactive multimedia in the light of the arguments developed so far, refining and developing the arguments while using them to make sense of the spatial properties of this medium, especially in so far as they relate to depiction. A taxonomy is proposed which is intended to reflect existing practice and indicate promising lines of enquiry for the future.

While the spatial practices of film have their own literature, those of interactive media, like those of television, have little. Given the relatively short history of pictorial interactive media, all categorisation – and analysis based on that categorisation – must be tentative.

In considering pictures, film, television and digital interactive media within a single coherent framework, it should become clear to what extent the spatiality of each may be fundamentally different from that of the others. In particular, as already indicated, it raises questions about the transferability of spatial practices from one medium to another, such as from narrative film to interactive media. The relatively new demands of interactivity will, I suggest, have a strong effect on spatial appearance, producing new developments and refinements of spatial practice. Genres which cannot be foreseen will come into existence and new spatial practices will be a defining characteristic of those genres.

In this chapter I will show that, like television, interactive digital media is not a genre but a technology supporting multiple (albeit emergent) genres and that the spatial practices required for each are likely to diverge as much as those documented in pictures and other screen-based media. While I have argued that the fiction film should be considered a mature medium, and the same is beginning to be true of factual television, I will suggest that the spatial practices of pictorial interactive multimedia currently mark it out as an *immature* medium in which form fails to articulate meaning. This is partly because there is insufficient understanding shared between makers and users, so that potentially expressive devices have not been assimilated into convention. I have made clear throughout the earlier parts of this thesis how the apparent transparency of spatial representations is in large part an outcome of the viewer's acculturation to the representational devices used.

Each genre of digital interactive media will, I argue, tend to become identified with one of the rival positions defined in this thesis, offering on the one hand an immersive sense of realism which is designed to seem as far as possible like unmediated *seeing* (even though significantly culturally determined), or on the other hand as a more objective mode of *showing*. The former is derived from the traditions of figurative pictures, the classical fiction film and the console game, while the latter resembles the overtly configurational media discussed in factual television.

Examples studied

Since anything which combines three or more media types – for example text, graphics and sound – in an interactive digital display may be classed as interactive multimedia, it is necessary to be selective for the purposes of this analysis. At one time considered a single genre of publishing associated with CD-ROM, interactive multimedia is now understood to take in websites, personal digital technologies, interactive television, virtual environments, point-of-information systems and many other distinct forms. Since my aim here is to investigate the relationship between the spatialities of pictures, film, television and interactive media, I incline towards those artefacts which, unlike for example a current WAP telephone, are able to support relatively high-resolution displays and in which the scope for spatial articulation is fairly considerable. The discussion concentrates on artefacts with significant pictorial components. Despite divergences in the form, content and context of interactive media artefacts it is possible to make some useful general observations on the basis of a few examples. Some of the artefacts discussed are unpublished projects which take a rather more adventurous approach to spatial articulation than their commercial counterparts; nevertheless, all are chosen to represent the kinds of multimedia currently deliverable on standard computers. This is partly because these were examples accessible to me at first hand for extended study and partly because, as I shall show, the limitations of the technology have forced the development of interesting spatial innovations which have not previously been analysed .

2 Some characteristics of spatial practice in interactive multimedia

Once it became graphical, interactive computing could not evade comparison with film and television, with which in some respects it struggles to compete. The book metaphor of early works has been invaded by, and in some cases replaced by, spaces derived from these fluid, image-rich technologies. In combining media which have their antecedents in mature spatial genres, the makers of interactive multimedia artefacts seek to imitate some of the apparent strengths of other media. This imitation may in some cases be counter-productive and take insufficient account of the differences in objectives and modes of use of varied artefacts. For example, pictorial interactive multimedia tends to seek the benefits of cinematography without having yet found a way to form these ‘shots’ into a coherent whole, mainly because there is no internal motivation which ties the events and therefore the spaces together as there is in narrative. I showed previously how the spatiality of the fiction film is not designed to exhibit pre-pictorial space but is essentially concerned with pictorial issues – it constructs just sufficient space to serve the objectives of storytelling and psychological engagement.

A characteristic of the screen display of interactive multimedia is that it is in a simple sense generally unrealistic: the issue discussed in relation to pictures and film as to how closely a visual representation, considered as a unitary display, can correspond to looking at real scenes seldom arises, principally for reasons of performance. Technical impediments to the imitation of scenes have a far greater impact on what may be done in multimedia on standard computers than in film and television: in fact

these limitations dominate the visual form. In the case of film and television to simply point a camera at a scene yields unlimited quantities of broadly realistic full-screen material, but this has not been the case with computer-based media, where not only has a storage medium such as CD-ROM been inadequate in both volume and speed of transfer, but standard computers have been unable to deliver fluent full-screen photographic sequences, especially when synchronised sound is also required. In the analysis of a number of interactive multimedia artefacts which follows, it will be seen that few even attempt the kinds of all-embracing realism associated with film and photography, though as with news television they may make extensive use of embedded mimetic segments, photographic or cinematographic.

Whereas in television the introduction of digital technology meant 'a new kind of magic' in which 'you can't see the joins'[1] its lower-powered equivalent in digital media has not generally succeeded in closely integrating discrete media components, and the early tools available did little to overcome this compartmentalisation. Despite the popular focus on 'convergence' – for example **Murray 1997 p27, Fischetti 2000** – in spatial terms the integration of different forms of content has proved elusive. Early popular development packages such as Apple's *HyperCard*[2] or Asymetrix's *Toolbook* treated each component as discrete, so that an item of digitised video footage would appear in a small rectangle resembling a self-contained cinema or television screen. More recently Macromedia *Director* when used in conjunction with image manipulation tools such as Adobe *Photoshop* has given developers greater facilities for integrating different media technologies in unified spaces, so that *QuickTime* movies, *QuickTime VR* scenes, still graphics and text can be combined moderately well into spatially and temporally integrated artefacts, though even with these tools such tasks are not trivial.

Where the technology permits it, the combining of pictorial and other segments in single displays is often taken still further than in factual television. Instead of segments being contained by rectangles they are frequently matted into one another to create a seamless surface, albeit one that partakes of multiple viewpoints, scales and even modes of representation. These amalgams of pictorial segments go beyond mere configuration of parts and approach the status of pictures in their own right, pictures whose spatiality is designed in response to the special demands of interaction.

Given the importance of the World Wide Web in terms of the quantity of artefacts produced and its prominence in public awareness, I should explain its general absence from this analysis. The structures available for the encoding and delivery of Web content are based on an even more strongly compartmentalised approach than in the early days of disc-based multimedia. Text is the only medium handled with

[1] Martin Lambie-Nairn in *Creative Review* September 1984, quoted in **Crook 1986 p11**

[2] HyperCard © 1987-1993 Apple Computer; Toolbook © 2000 Click2Learn.com Inc.; Director ©1985-1998 Macromedia Inc; Photoshop © 1989-1996 Adobe Systems Inc; QuickTime © 1989-1999 Apple Computer; QuickTime VR ©1991-99 Apple Computer; Shockwave © 1985-1998 Macromedia Inc; Shockwave Flash © 1986-1999 Macromedia Inc.

any facility by HTML[3] and its derivatives such as CSS,[4] all other media being little more than an afterthought. Early Web browsers even on graphical computers displayed nothing but text, passing over the display of non-textual segments to ancillary software which displayed them in a separate window. Gradually the use of tools such as Macromedia *Shockwave* and *Flash* is enabling greater visual and temporal integration simply by arrogating control of the browser window (Figure 7.01), but at the time of writing these are of minimal importance for the majority of Web users (**Nielsen 2000b**). Nevertheless, I assume for the future a far greater range of spatial usages made possible by such technologies, and therefore concentrate on them at the expense of standard Web documents of today.

Figure 7.01.
Good Technology 2000: A website for the band U2.
The site makes extensive use of modes not possible using the standard interface devices available in HTML. In particular the interactive timeline in the lower half of the display allows visual properties such as translucence and interactions such as dragging.

3 The importance of interactivity

Interactivity is not just an additional layer imposed over existing forms of spatial articulation but fundamentally influences how those forms are deployed. Text, easily manipulated by the computer, has become capable of change in spatial configuration on demand;[5] however this is rarely the case with pictorial or filmic material. Generally each segment is of fixed form and even duration, though occasionally the user may be permitted to resize a pictorial or filmic component (Figures 7.02a-b).

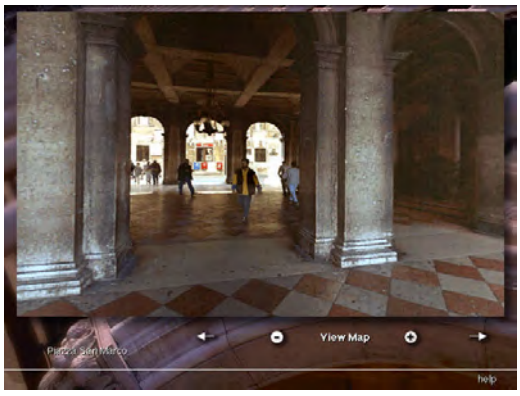


Figure 7.02a-b. Editoriale Domus: CD-ROM: *Venezia*, 1998.
The user may switch between two sizes for the presentation of QuicktimeVR scenes, principally in order to overcome limitations of performance. The presence of the large VR segment leaves no space to display the map and some of the controls. In both cases a photographic background image fills the space not occupied by the main segments. Note also how objects are embellished with drop-shadows (such forms of 'realism' were discussed in Chapter 4).

Nevertheless, even for pictorial and filmic material, basic differences are introduced by interactivity. A decisive influence is the fact that the user must be provided with

[3] The Hypertext Markup Language, which specifies the syntax of a Web page for interpretation by a browser.
[4] Cascading Style Sheets, an extension to HTML which give somewhat greater control over the positioning and layering of graphical and textual components in Web pages (**Lie and Bos 1997**).
[5] For example in outliners, diagramming software, presentation packages, database visualisation, hypertext systems, file management systems and so forth (discussed extensively in **Boyd Davis 2000**).

objects with which to interact by pointing, normally by directing a cursor using a mouse, trackpad or similar device. Pointing requires the presence of an object on the screen at the time when the user wishes to make the interaction. It is thus quite unlike an object in the space of film or television which has earlier been seen by the viewer and which is assumed to be still present even when not on screen. Though Poole (**2000, p73**) lumps together joysticks, joypads, mice and keyboards as 'curiously alienating devices' it is pointer-based interaction which dramatically constrains spatiality.

As well as marking a difference from film, the need for objects to be visible in order to be interacted with is one of the principal ways in which the graphical user interface is different from, and inferior to, the command line interface which preceded it. There a user could type **copy a:myfile b:** to move a file from one volume to another without any need to see a prior representation of the objects referred to. Multimodal interaction, allowing the user to interact using a mode such as speech (**Cassell *et al* 1999**, **Oviatt and Cohen 2000**), could radically affect this spatial characteristic of interactive multimedia since users would then be able to address objects which they could not see.[6] Though Grasso, Ebert and Finin (**1998**) enumerate several points of comparison between graphical and multimodal user interfaces, they omit this basic fact, that graphical interfaces are bound to make all available objects visible. This is a fundamental problem for most of the artefacts under discussion.

If many objects must be visible at once, an important determinant of the spatial characteristics of pictorial interactive multimedia is the issue of information density. Many characteristic spatial features arise from this difficulty. Using 9-point text on a display 800 by 600 pixels, about 350 textual labels or titles (approximately 740 words) might be presented in a table. Using continuous text, the number of words which may be displayed increases to about 1,500. Using some method of indirection, where for example each item is represented by a small dot with which the user interacts in order to access the information itself, it might be possible to display as many as 13,400 items, assuming 3x3 pixels to represent each point and sufficient space between points to distinguish them. Pictorial and filmic segments however are considerably larger than their textual or schematic 'equivalents'. I suggest it is hardly possible to make a picture which is usable for most purposes comprising less than about 7,000 pixels. Fewer than seventy such pictures can be fitted into an 800 x 600 display, so pictures pose a considerable problem of density and space consumption.[7] An inevitable outcome is that designers must often choose when to distribute pictorial elements spatially and when to use time as a substitute for space – for example by creating a pre-pictorial space of which the user views a portion at any one time.

While the need for the objects of interaction to be visible is one aspect of the special

[6] Sound has the potential to deal with another aspect of the same problem, if acceptable ways could be found to have objects currently outside the display emit sounds which reminded the user of their presence, thus replacing the binary cut-off of the screen frame with an analogue decay with distance.

[7] Almost 470 icons of 1024 pixels (32 x 32) can be fitted in such a display but only if packed edge to edge, and in any case such pictures are generally useful principally as visual mnemonics rather than true depictions. They have the realism of general recognition but almost no other. A 'passport photograph' portrait may be considered reasonably recognisable as a particular individual at 70 x 100 pixels.

demands of interactivity, another is the relationship between narrative and interaction which I have already suggested makes difficult any direct transfer of filmic practice to interactive multimedia. As examples will illustrate, a basic problem of the spatial design of interactive multimedia is that the more freedom is given to the user the greater the difficulty of using the practices of spatial articulation exploited in film, since these, as I have demonstrated, are outcomes of authorial control. I attempted in a book chapter on virtual environments (**Boyd Davis and Athoussaki 1999**) to discern ways in which the spatial techniques of cinema could be repurposed in interactive systems, especially through greater use of the variables of the view such as selective focussing, but I now recognise more clearly the difficulties of such borrowings. I also underestimated the importance of the illusion of unmediated viewing. Persson (**1998**) similarly proposes ways in which the spatial practices of film might be adapted to support users in making sense of virtual spaces, for example in order to assist them in finding their way about. The difficulties of such transfer are discussed in relation to particular artefacts below.

4 A spatial classification of pictorial interactive media

In what follows I categorise the uses of space according to principles arising from the analysis of pictures and of film, proposing six categories:

1 Simple assembly
2 Two-dimensional pre-pictorial space
3 Three-dimensional pre-pictorial space
4 Pseudofilmic space
5 Hybrid space
6 Integrated spaces combining pre-pictorial and pictorial space
7 Pictorially dominated space

These range from simple assembly of pictorial components in the plane and in (shallow) depth, through to what I suggest are new pictorial forms specifically designed to support the demands of interaction and the limitations of the technology. Not surprisingly these new forms are largely derived from inherited practices reinterpreted in relation to the new objectives. Other categories bring out the different kinds of relations between the design of the pre-pictorial model and of the pictorial view by which it is mediated. This will be seen to be correlated with the mode of interactivity – allowing the user substantial freedom of viewpoint diminishes pictorial expressivity. Various designers have wanted to keep hold of the expressivity granted by authorial control of the view and have as a result concentrated interactivity into other aspects. The analysis concludes with some suggestions of ways in which the spatial representations of pictorial interactive multimedia will diverge according to the variety of objectives.

1 Spatial configuration: simple assembly

Under this category I include discussion of spatial organisation which is schematic rather than pictorial, since its inclusion will facilitate later discussion of systems having a stronger pictorial component.

The issue of information density, arising from the need to bring together multiple elements in a single space in order to make them all accessible not only to vision but also to interaction, leads in some cases to segmented spaces somewhat akin to those of the television news broadcast, using straightforward juxtaposition and layering of related elements. Like the news display, this sometimes approaches a kind of pictorial unity, but more often consists of a segmented configuration in which the components are arranged on simple principles such as the matrix, or by a depthwise accumulation of segments using basic occlusion. In the spatial organisation of interactive tools, such as applications like Adobe Photoshop or Macromedia Dreamweaver, this has principally taken the form of an accretion of numerous more or less standard interface objects such as windows, palettes and tool bars (Figure 7.03a). The screen furniture of these components serves to divide up the display into discrete areas. The relationship between them is not articulated visually, so that it is not generally apparent which objects control which until the user has learned the working methods for each particular application.

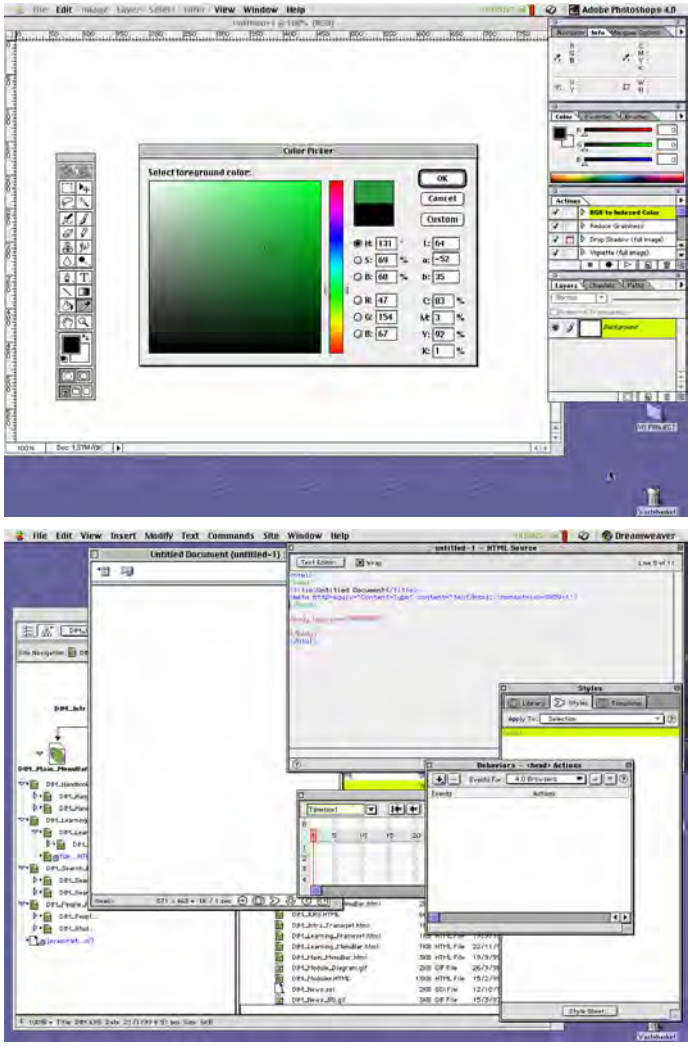


Figure 7.03a-b.

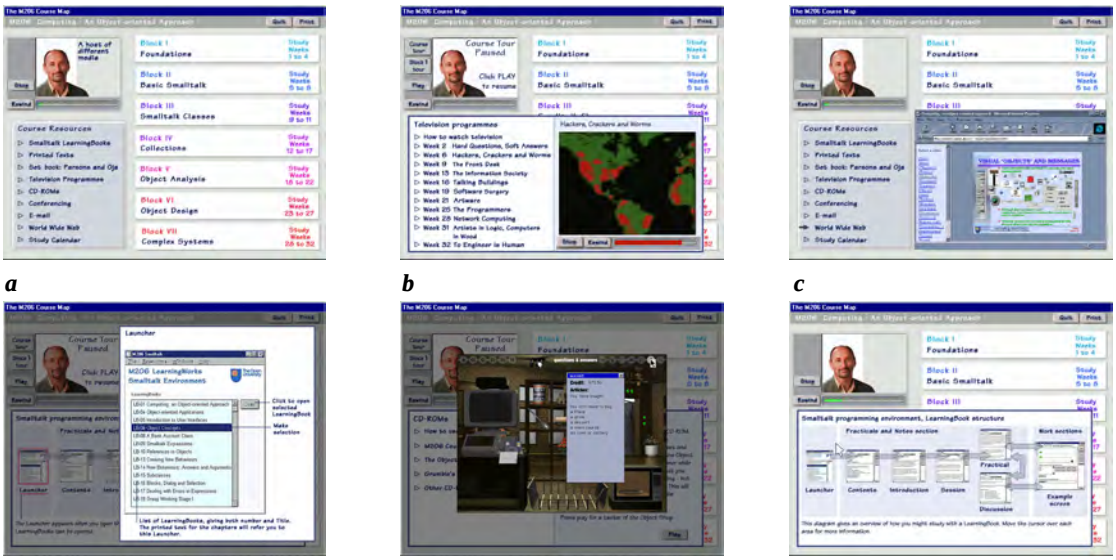
a: Adobe Photoshop Version 4.0 ©1989-1996.

b: Macromedia Dreamweaver Version 2.0 ©1997-1999.

The multiple windows and palettes typical of complex computer applications.

Figures 7.04a-f illustrate an application of a slightly different kind. Intended for those new to computing, the *Interactive Course Map M206* (**Open University 1997**) makes use of overlaid as well as juxtaposed components but imposes a stronger authorial configuration on them than do the tools for advanced users just described. The user

may open and close overlaid panels but not move them. This artefact lies at the intersection of the authorially controlled and the fully interactive, offering a multimedia presentation with a strong narrative thread as well as a more interactive, open information structure which the user can interrogate at will. The user may follow a voiced narrative which explains each aspect of the subject matter (the syllabus and educational resources of a course in object-oriented programming) or click on various segments of the interface to access hidden detail. The design borrows from the strongly delineated discrete components of application tools but articulates



a ***b*** ***c***

d ***e*** ***f***

Figure 7.04a-f.
Interactive Course Map for M206, Open University 1997.
The application uses overlaid components but incorporates them into a more strongly configured whole.

the relationship between them rather more clearly using spatial configuration both in the plane and in depthwise layers. For example the imposition of spatial cues of figure and ground (Figures 7.04*d* and *e*) makes clear that only the foreground components are accessible. In Figure 7.05 an interesting comparison emerges between the strongly delineated components of the tool itself and the merged spaces and modes of representation of the embedded television sequence which combines presenter and animated map in a homogenised and ambiguous space. While both the television programme and the multimedia course map make use of segmented spaces they do so in different ways, determined partly by their respective antecedents and partly by what the technology of each makes easy. Considered as a whole, the program presents a multiplicity of variations on the idea of combining planar segments into meaningful configurations while preserving their discreteness.

One of the simplest devices used in the *Course Map*, as it is in standard computer interfaces, is the opaque overlay which occludes the underlying material. This is one of the spatial usages based on depth-wise stacking which is badly managed in the standard computer interface. The lack of visual semantic relations between the windows in the accumulated pile in such interfaces, together with the simple fact that the windows which the user needs access to are often obscured by others, has been

Figure 7.05
A multiplicity of segments can be seen, each delineated as discrete except in the case of the blended spaces and modes of representation visible in the embedded television sequence.

accused of leading to poor performance by users when compared with more fully articulated configurations which preserve the visibility of every item (**Kandogan and Shneiderman 1997**). A related typical weakness is a lack of 'intelligence' in the positioning of windows so that for example a *Find* dialogue-box obscures part of the document in which searching is taking place and, though the underlying document scrolls to reveal each successive occurrence of the search term, it is not 'aware' of the presence of the occluding overlay so that occurrences are highlighted invisibly beneath it. The designers of the *Course Map* avoid the use of movable overlays and so

Figure 7.06a-b. **Dorling Kindersley 1997**: Sampler of Dorling Kindersley multimedia products. Typically a pop-up overlay obscures the object to which it refers.

aggravate the problem of occlusion since offending objects cannot be dragged out of the way, but in compensation they have done their best to position overlays where they will cause least difficulty and where they provide useful glimpses of the underlying context. The same cannot be said of the CD-ROM products of Dorling Kindersley which invariably position overlays in the centre of the display, the very location where the focus of interest in the underlying screen is itself likely to be located (Figure 7.06a-b). Frequently additional pop-up overlays can be summoned over one another so that the original subject to which they refer seems in danger of disappearing under an accumulation of boxes. This depth-wise spatial configuration is expressive of the basic fact that the new information is ancillary and that it may be dismissed by the user, but neither it nor the planar configuration contribute anything else to the meaning.

Figure 7.07.
Donald Norman 1994: *Donald A Norman – defending human attributes in the age of the machine*, CD-ROM 1994.
Pictorial space contained more or less conventionally within a page-like assembly.

A complementary example of combined text and graphical space is offered by a Voyager CD-ROM based on several works of Donald Norman (**Norman 1994**). It works effectively as a 'book' with the additional digital benefits of free-text searching and facilities to easily mark and index passages for future reference. The illustrations function in most respects like those of a traditional book but with the ability to animate and use sound (Figure 7.07 – I comment later on the characteristics of the graphic segment in this figure when considered in its own right). The physical relation of picture and text is one of simple juxtaposition.

Figure 7.08a-b. **Maltez, Bennett and Cova 1997**: Interactive Documentary 'Contact'.
In ***a*** the main televisual narrative sequence alone is available. In ***b*** the main narrative is partly suppressed when the user has activated one of the ancillary narratives which is offered at a time determined by the authors.

Whereas the spatial articulation of the interactive course map was a development of spatial usages familiar from application tools, *Contact* (**Maltez, Bennett and Cova 1997**), a prototype interactive documentary, adopts an approach which slightly more resembles television (Figure 7.08a-b). The task which it addresses is less ambitious than that of the course map, yet there are more points of similarity than their respective interfaces might suggest. Both offer a continuous narrative in which every component is at some time embedded; in both narratives the user may advance or step back by interacting; both offer ancillary information at predetermined points in the overall narrative; and both allow the user to interrogate individual components of the interface. However, in *Contact* the user has not the same degree of arbitrary access to all the components and the sensation of authorial control is correspondingly stronger. The number of individual components in *Contact*, and the range of relationships between them, is smaller than in the *Course Map*, which makes easier the provision of a simple user interface.

The spatial differences between the two artefacts are evident in the illustrations provided here, but the differences are augmented by the style of interaction. While the *Course Map* is based on the now conventionalised screen furniture of standard graphical user interfaces which are often considered examples of *direct manipulation* (**Shneiderman 1992 p202-205**), the form of interaction offered by *Contact* is often *more* direct in that it involves manipulation of the objects of interest rather than with control devices which in turn affect objects. For example, to control a digitised video in the *Course Map* the user drags a slider, while in *Contact* the user drags the pointer to right or left on the image itself. In the *Course Map* the user activates an element in the display by clicking the pointer on it – that is, by depressing and releasing the mouse-button – while in Contact the user merely moves the pointer into the segment of interest. This minimal action is sufficient to cause the indicated segment to become more prominent and to perform (it brightens and its narrative proceeds) and to suppress the other narrative, which is dimmed while its narrative is suspended. The relative seamlessness of the visual interface is therefore enhanced by the mode of interaction.

The contrast between controlling video in the *Course Map* and *Contact* illustrates the problem of *maturity* (or rather the lack of it) in interactive media. At present the use of a slider and buttons (a hybrid of the scrolling window and the video tape recorder) has become familiar, while the direct control of video by interacting with its image has not. It is easy to imagine a situation in which the reverse is true: the provision of external controls may come to seem obscure and alienating and the direct style natural and intuitive.[8] Only familiarity, the sharing of a convention by makers and users, could make this happen. The simplified presentation and direct style of interaction taken together give some of the sense of unmediated access to content which I noted in relation to film, and emphasise the need to conceive of visual and interaction design as aspects of a single set of objectives.

Despite their differences *Course Map* and *Contact* both preserve the discreteness of segments. A different approach is taken in *Britain in Brief*, a CD-ROM for the Foreign and Commonwealth Office (**1997**) (Figures 7.09 and 7.10a-c). Here there is an attempt to bring multiple images, each with its own perspective geometry, together in a

Figure 7.09.
Foreign and Commonwealth Office: *Britain in Brief*, 1997.
A digital collage of multiple mimetic images. Each object or scene has its own perspectival geometry, but the overall tonality is adjusted to equalise the salience of each of the objects.
In this illustration a pop-up text is also displayed (below centre, left) which seems to emerge out of the overall assemblage. Considerable effort has been made to make this text box soft-edged and translucent.

[8] The advantages of indirect manipulation discussed earlier in controlling non-visible objects cannot be claimed for the slider and buttons, since it is never likely that the controls will be visible in the absence of a filmic sequence itself.

a

b

c

Figure 7.10a-c.
Foreign and Commonwealth Office: *Britain in Brief*, 1997.
Figures ***a*** and ***b*** show two states of the same screen. In ***b*** a small digital filmic sequence (using QuickTime) is displayed in place of a static graphic. Figure ***c*** shows the post-processing applied to each frame of the QuickTime movie in order that it should appear integrated in the overall collage-space, disguising the fact that the movie is in fact rectangular.

digitally collaged space. The motivation is presumably similar to that behind the blending and spatially ambiguous techniques of television – to suggest an indefinite wealth of information (as well as perhaps simply being the fashion of the moment). The visual blending on which considerable effort has been expended begins to offer a quasi-pictorial space complete with relations of scale and depth between the parts, but I suggest this is more of a graphic conceit than a meaningful expression of relations between the parts, which are essentially just accumulated.

2 Two-dimensional pre-pictorial spaces

Figure 7.11 (left).
Foreign and Commonwealth Office: *Britain in Brief*, 1997.
The display offers a partial view of a large planar space containing symbols for thirty topics such as Parliament, Tourism and Religion and Beliefs.

Britain in Brief has the characteristic common to many multimedia artefacts that its various parts use different approaches to spatial organisation. In addition to the browsable digital assemblages just described, and a televisual opening sequence, another interface offers a six-by-five array of iconic pictures based on photographs. Clicking one of the pictures leads to one of the themed digital collages already described, while simply pointing at it causes the display of a label, the production of an appropriate sound, and a small animation of the picture. The pictures themselves are of some interest spatially, in that, in common with many components of the computer interface, they cavalierly combine spatial realisms derived from being photographic with the contrived realism of objectness imparted by cast shadows. This

pragmatic attitude to combining the different realisms associated with different modes of representation will be seen to be more fully exploited when other categories of interactive spatiality are analysed below.

The user is never permitted to see all thirty pictures at once since only about four fit the display (Figure 7.11) and there is no facility to zoom-out for a broader view (nor, as a Web-user might expect, to alter the frame of the window). This is clearly deliberate, since there would be sufficient space to exhibit thirty such images at a usable size in a single screen – the designers have obliged users to scroll for some reason. There are no tools such as scroll bars for effecting the movement of the view in relation to the space, nor need the user drag (that is, move the mouse while holding the mouse-button down): the space moves if the user simply moves the pointer towards the perimeter of the display. As in the modeless activation of alternative narratives in *Contact* this has the effect of minimising the degree of conscious motor control of the interface and has some of the qualities of simply shifting one's attention. The nearer the pointer is to the perimeter the faster the display moves, avoiding the sense of transitional time wasted which I noted when a camera moves under authorial control from one significant part of a scene to another.

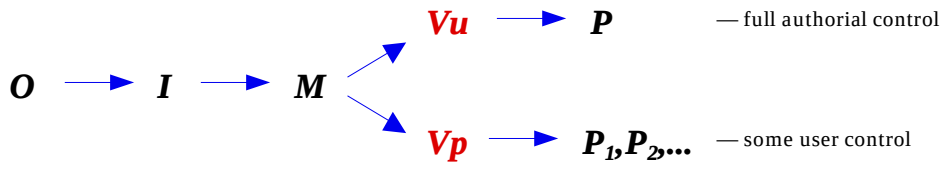
I define this scrolling surface as pre-pictorial in the sense that the scrollable layer can be considered to belong to a model, albeit itself containing pictures, on which the user may choose any view, and in which these views are not responsive to what is viewed. Like an imaginary camera (not a real one) the window renders in unmediated fashion the view of the model beyond. No aspect of the view is altered in response to what is viewed, unlike in painting for example where I showed that it is normal for a picture to re-present pragmatically and constructively the observed scene – to *create* the space of the scene – rather than simply to open a frame upon it. I also showed how film too, despite exploiting the mechanised automatic picture-making of photography, is essentially pictorial because it designs the model and the view as aspects of each other, manipulates the variables of the view and *makes* space through the relations between discrete shots. Here by contrast the window really is just a window.

Returning to the model proposed previously, a further refinement is required. I proposed previously that the the *objectives*, non-visual *idea*, pre-pictorial *model*, and *view* could be crudely diagrammed in relation to the *picture* as follows:

O → ***I*** → ***M*** → ***V*** → ***P***

In the case of a still picture, film or television programme, the author has control over the specific view of any pre-pictorial model, which offers a particular picture. When considering interactive systems, however, the user often has the freedom to choose which parts of the model to view, from what distance and what angle. Nevertheless it would be quite untrue to say that the user has access to the model itself. As defined, the model is not a visual entity (though its form does imply some visual characteristics). As already stated, the model is given visual form by the parameters

of viewing and depiction, and in almost every case these are substantially decided by the author, not the user, imposing strong limitations on the kinds of pictures which may be seen. This may be represented as follows:



Here the use of pictorial means under full authorial control is represented by the upper branch, in which uniquely prescribed views ***Vu*** yield particular pictures ***P*** over which the user has no control. The lower branch represents those interactive systems in which viewing parameters ***Vp*** rather than actual views are decided by the author. These yield an unlimited number of pictures $P_1, P_2, \ldots$.

The Britain in Brief scrolling interface shows some odd characteristics in the relation between its pre-pictorial and pictorial aspects. The user might expect the boundary of the scrolling model-space to soon be reached, as when scrolling pages in a word-processor or a Web browser. However this is not so since attempts to scroll continuously either horizontally or vertically eventually show the whole space repeatedly: the space turns out to wrap back on itself. This infinite topography could be conceptualised as a torus within which the user is located, but there is no visual evidence of curvature – everything about the graphic suggests that the model is a plane. As with film, it is experience of the visuals over time which constructs this 'space', as much as any single view of it, but in this case the single view – suggesting a plane surface – and the aggregation of views – suggesting a toroidal model – offer contradictory evidence.

Such indeterminate or contradictory spaces are potentially a characteristic of any non-physical medium but digital methods facilitate them, and there are clear precedents in the predominantly textual realm of databases and similar digital structures where objects can be in more than one space and more than one configuration at once. If a new form of non-physical space proves on repeated and extensive use to offer advantages in relation to the objectives of the artefacts it serves, no doubt conventionalised nomenclature will emerge to normalise these unfamiliar concepts and they will become transparent as so many filmic techniques have done.[9]

The intention of the authors of this particular space, declared on the packaging, is to offer a 'voyage of discovery' in which users can 'explore.' Another probable motive was to give the impression of an unending wealth of information. In both respects the product seems likely to disappoint since it soon becomes clear that there is a limited number of items, while at the same time it is not easy to find any particular one, even if it has been seen before, because of the featureless character of the background. Nevertheless such a spatial system, redesigned for more straightforward use, has the

[9] Similarly the bizarre metaphoricity of 'scrolling a window on the desktop' has decayed to simple naming of now familiar devices and behaviours.

potential to solve some of the problems of information density while supporting a fluid interaction style which obtrudes only slightly into the process of looking.

While the space described supports simple browsing it is clearly incapable of supporting in its present form other functions such as searching. The need to support multiple means of access to information – something that interactive *textual* configurations are generally good at – is part of the rationale for the provision of more than one graphical organisation in a single artefact. A particularly poor example, which I include as a measure of the immaturity of the spatiality of interactive multimedia, is a recent companion CD-ROM to the BBC's *A History of Britain* television series (**V&A/BBC 2000**). Like many this product offers more than one spatial organisation of the material yet it is unclear why. Many possible browsing and searching functions which might have been offered are not, while the two interfaces which are provided (Figure 7.12a-b) duplicate one another: the 'timeline' and 'stories' interface differ only in that one must be scrolled and is horizontal and the other need not be scrolled and is vertical. In both screens the disposition of objects in one axis very approximately represents time while that against the other axis is meaningless. Poor design exists in all media; the significance here is that it would be hard to imagine organisations such as those involved in the production of this CD-ROM – the BBC, BBC History Magazine and the Victoria and Albert Museum – sanctioning work of such incompetence in a medium with which they were more familiar. There is a strong impression of designers failing to find any fit between the new spaces made possible by digital interactive media and the functions which the artefacts are intended to perform, and little suggestion as yet of a mature spatiality which gives form to meaning, which is expressive in the way that the customary spatial practices of older media are.

Figure 7.12a-b. **V&A/BBC 2000:** *A History of Britain* CD-ROM.
The 'timeline' and 'stories' interfaces differ only in insignificant ways.

For new spatial forms to become adopted into a developing spatial 'language', use must reveal a close relationship between the spatial form and the makers' intentions. When what is intended by the provision is unclear, spaces are doubly incomprehensible. The user has the problem not only of understanding an unfamiliar spatial device, but of attempting to discern an intention behind it. Whereas spatial innovations in film, such as the close-up, generally had discernible motivation, this is

often not the case in interactive media.[10] I noted earlier how film-makers tended to abandon unmotivated devices in favour of those which articulated meaning. I do not wish to suggest that innovation derived from open-ended experimentation with technical possibilities is to be deprecated, but it seems clear that over time only those devices which in context have or acquire meaning will flourish.

3 Three-dimensional pre-pictorial space

In the case of the three-dimensional world presented in the plane of pictures and film, every aspect of the artefact may be regarded as pictorial, since the depicted model is chosen and designed with a view to appearing in particular views, views whose variables are chosen by the maker and which in the case of film are also sequenced with a view to articulating particular meanings. In the case of pictorial interactive multimedia the situation is more complex. Some of the example three-dimensional spaces discussed here are best considered as pre-pictorial. Users can look where they wish in the space and no accommodation to the model is made in the view – the view is an entirely automatic outcome (like cinematography but lacking any deliberate interventions) of the virtual viewing device's processing of the model. Other examples, discussed next as *pseudofilmic spaces*, are more truly pictorial in a similar sense to film in that, though the user may move about the space to a certain extent, every view has been chosen by the maker. The users' decisions as to which trajectory to follow are responded to by the showing of pre-rendered views. Though the user has some freedom of movement, the mode of viewing largely resembles the *optimal view* discussed in relation to film.

The use of sequences of separate views of a three-dimensional model is in many cases a consequence of limitations in delivery technologies. Since, as discussed, standard computers have not been well suited to the provision of smooth-flowing filmic sequences rendered on the fly as users move their viewpoint, the pre-rendering of selected views has presented a solution to this practical problem. This might be regarded as a disappointing compromise, but I suggest that some of its benefits may not be abandoned even when technological advances make it unnecessary.

To the class of pre-pictorial spaces which allow the user largely unconstrained and unarticulated views of the model belong most *virtual worlds*, such as Active Worlds (Activeworlds.com Inc.) in which users build three-dimensional environments which can be viewed in a suitable browser (Figure 7.13.). It is certain that the geometry and other attributes of this world exist independently of any particular view: a given depiction at any moment is the automatic outcome of the application of generalised viewing parameters to the world-data, so there are none of the opportunities to engineer a particular pictorial outcome in order to fulfil particular objectives which are so important in authored picture-making. Nevertheless, users may not really look entirely where and how they please in such environments. For example they are

[10] The presence of a rationale for technical innovation could be overstated for film. 'What mattered to me was a fine close-up. It so happened that if they were to accept a close-up the public had to be given a story' (**Renoir 1974 p56**). Nevertheless mainstream film-makers tended over time to eschew gratuitous effects and exploit technique in the interests of narrative.

never permitted to reach the pictorial horizon – it is infinitely far away and no amount of travelling will take the user to it. The focal length of the virtual lens is fixed and certain viewing angles are forbidden; for example, the viewpoint may not pass in an arc through the vertical, probably to avoid the disorientating inversion noted in the chapter on Film (or perhaps to protect the program from divide-by-zero errors). In addition, since the user is by definition an inhabitant of such shared spaces, the viewpoint is always either through the user's eyes or over the avatar's shoulder (a clear borrowing from film).

Research with young children in Active Worlds by Bailey and Moar (**2000**) casts further light on expectations of filmic convention. Children who were enabled to apply scanned photographs of their own faces to their avatars wanted to be able to see not only the faces of their friends but their own as well. Such views of the protagonist would, with the rare exception of experimental first-person films like *The Lady in the Lake* (**Montgomery 1946**), be a standard part of any film and it is tempting to think that this experience of another, powerful screen-based genre, rather than the experience of real life, prompted the children's desire. An *ad hoc* solution sufficient to satisfy the children was to apply the face to the back of the avatar's head as well as to the front so that it was always visible in the over-the-shoulder view, a nice indication of the selective approach to realism which the conjunction of an unfamiliar technology and a set of objectives can engender.

Figure 7.13.
Active Worlds: Scene in Active Worlds 2.2.
© 1995-2000 Activeworlds.com, Inc.
In this Web-based virtual environment, there is no doubt of the distinction between the pre-pictorial (three-dimensional) space and the (two-dimensional) pictorial space of the display.

In Active Worlds, the space itself is a substantive part of the experience, rather than being simply a means to an end, but three-dimensional environments are increasingly used to offer access to data-sets, where the information content is the substantive feature.

Kullberg has developed a three-dimensional timeline (Figure 7.14). An interesting difficulty arises in its attempt to support multiple modes of access within a single space. The user may navigate the data by 'travelling' through the space, but also using 'virtual' routes which connect sequences of data-items in non-topographic ways, for example through the results of a search operation. In this mode, time is used in place of space to present a sequence of face-on views of each item, but they

still have their visual framework of adjoining items (Figure 7.15). The surrounding information which in smoothly animated 'fly-throughs' of the space enhances the sense of concreteness and imparts useful context is now experienced as a distraction.[11]

Figure 7.14.
Robin Kullberg: *Dynamic Timelines: Visualising Historical Information in Three Dimensions*, 1995.
From **Kullberg 1995**.

This highlights the problems of constructing pre-pictorial spaces which fail to support all the modes of interaction demanded of them. Although no mimesis of the world is involved there is nevertheless an excessive adherence to the unmediated viewing of a pre-pictorial three-dimensional structure. The example of film suggests that the mode of picturing should not be independent of the objectives in this way. In a close-up in the classical fiction film the viewing parameters would be quite different – probably the background would be both darkened and defocussed – in order to support the different kind of viewing relevant at that juncture: the unmediated depiction of the pre-pictorial model fails to provide such an optimal view.

Figure 7.15.
Robin Kullberg: *Dynamic Timelines: Visualising Historical Information in Three Dimensions*, 1995.
The user can jump through a series of images without navigating the space. Contextual information which was useful when navigating becomes distracting when experienced outside the navigational experience.
From **Kullberg 1995**.

4 Pseudofilmic space

Within broad limits in the systems just described the user is allowed to look anywhere. A related class of artefacts also offers full-screen views of a pre-pictorial model, but this is explicitly mediated by the maker since only views which have been pre-rendered are available. The user infers that there is a single consistent pre-pictorial model of which these views are scenes, in the manner familiar from film.

In each of the artefacts discussed now, some degree of entertainment is intended. Examples include CD-ROM 'edutainment' titles and adventure and strategy games. In all cases there seems to be an attempt to make the space immersive by eliminating counter-cues such as the paraphernalia of the everyday computer screen, or at least to

[11] My observations of the Kullberg project are based on a 5-minute video and the author's MSc dissertation (**Kullberg 1995**). Attempts to engage in correspondence with the author have been unsuccessful.

subordinate such items as far as possible. In this they aspire to the apparently unmediated qualities of the fiction film. The success or failure with which this quasi-cinematic usage is implemented is instructive.

Figure 7.16a-c.
Dawkins 1996: *The Evolution of Life with Richard Dawkins.*
A borrowing from film: the photocopier provides a distinctive point of reference which allows the user to infer continuity between three angles of view. However there are weaknesses in the way the technique is applied.

In the case of *Evolution of Life* (**Dawkins 1996**) there are two principal 'sets', one of which is apparently Dawkins' Oxford study and the other is a slightly futuristic gallery space. Some spatial practices used are borrowings from film. For example the prominence of an identifiable object in multiple views helps users to infer that these are aspects of the same space (Figure 7.16a-c). As Persson points out (**1998**), object-continuity in film is generally supported by the familiarity of objects and scenes depicted, which make it easy for viewers to surmise the parts they cannot see and to construct a whole environment from a series of glimpses.[12] In Dawkins' study there is a set of views which users see when they choose to rotate their viewpoint. These consecutive views are orthogonal to one another, which has just the tendency to draw attention to the medium which the fiction film studiously avoids (and Greenaway revels in). This is inappropriate here where it seems that there is an intention to make viewing seem naturalistic. The cutting from one shot to another is not in itself a problem and is as I have pointed out standard practice in film, but the mature art of 'transparent' film editing would have ensured that the distinctive item whose object-continuity is used to tie consecutive shots into a single space (here the photocopier) was matched for position from one shot to the next, whereas it jumps to the opposite side of the frame (Figure 7.16a and b). Insufficient attention has been given to the pictorial aspect of this quasi-filmic practice when compared with the care lavished on the pre-pictorial model.

The user begins to be accustomed to looking into the three-dimensional space despite the awkward means of moving within it. However shortly afterwards, on attempting

[12] This is one possible argument for the situating of abstract data in realistic objects which is additional to the usual arguments for metaphor in the interface (eg. **Andersen 1990 p155**, **Preece *et al* 1994 p456-461**, **Nardi and Zarmer 1993**).

to pass through a door by clicking on it (a simple form of direct interaction with a diegetic object), the user is suddenly presented with a textual explanation which appears in the immediate foreground and suspended in the air (if it can be conceived as *in* the space at all). It is a strong and sudden reminder of the planar nature of the display as a whole, undermining the preceding attempts to make the screen transparent to a pre-pictorial world. Again it would be wrong to object that the difficulty is caused by a lack of realism. That would be to invoke a simple, unitary concept of realism which I have been at pains to demolish. The problem here is not one of realism as such but that two incompatible spatial modes have been combined. The floating notice (Figure 7.17) belongs to the objectifying spatiality of the Dorling Kindersley pop-up window (though at least it does not occlude the object to which it refers!) and effectively destroys the newly created immersive space of the pre-pictorial set.

Figure 7.17.
Dawkins 1996: *The Evolution of Life with Richard Dawkins.*
Awkward intrusion of other spatial modes: when the user clicks on the door at the back of the room to gain access to other spaces, a pop-up window appears in, or perhaps over, the space.

In *Art of Singing*, from the same company (**Notting Hill/JHM 1996**), some of the difficulties raised by disconcerting switching between orthogonal views have been avoided. More care has been taken to sequence views so that they seem to be part of a trajectory, though there are still times when consecutive shots fail to suggest continuity of location and the user must make a conscious effort to integrate two views, which predictably breaks the sense of immersion. The standard optimal view principle of showing only the relevant is used to advantage. For example, when users choose to get into the lift that faces them, the next thing they see is the view normally taken by someone riding in a lift – the view back towards the door – whereas in a virtual environment users would be obliged to rotate their own point of view, with all the inclusion of irrelevant visual information discussed in relation to such shots in continuous camerawork.

Figure 7.18.
Notting Hill/JHM 1996: *The Art of Singing*
One of the pre-rendered views in a trajectory through the building. Shots are post-processed to increase their painterly qualities and distinguish them from the photoreal.
A portable computer (bottom-right) remains in the periphery with viewers as they travel.

Though clearly intended to evoke a kind of pre-pictorial space, *Art of Singing* employs some painterly effects which emphasise that these scenes are mediated with artistic and metaphorical intent, rather than being transparent photo-realistic shots of a real place.[13] This kind of mediation seems to fit quite naturally with the objectives of the artefact.

In this same CD-ROM, there is an attempt to give the user a means of gathering information encountered while touring the virtual building. On entering the space the user is 'given' a portable device and subsequently this is present in the periphery of the screen whenever there are texts available that the user may wish to capture (Figure 7.18). The floating device exists in a rather awkward visual relation to the main scene, but it does properly support the function to which it pretends. If the main scene were presented as a fluid sequence of moving images rather than as a series of 'shots' then the peripheral portable would appear more strongly attached to the user's own frame of reference in a manner analogous to those floating screen elements in the corners of the factual television display described earlier. This permanent attachment of objects to the periphery of the display seems more effective than when such elements pop up suddenly in the diegetic space, and is common in the design of many games (see Hybrid Spaces, *next*). It is also a situation familiar to the car-driver whose view is constantly accompanied by items such as the tax-disc on the windscreen.

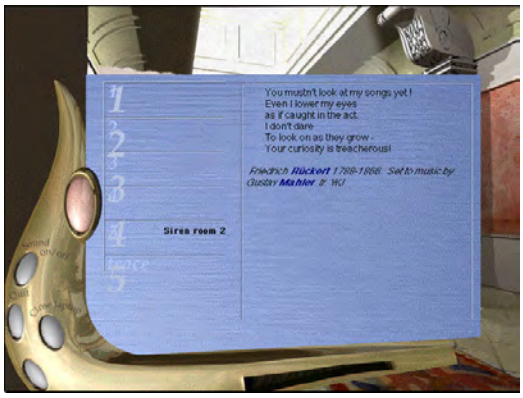


Figure 7.19.
Notting Hill/JHM 1996: *The Art of Singing*
The 'portable console' in which users can collect and view texts.

The texts which may be collected and viewed using the device are displayed in a modified version of standard text fields but these are situated in the portable console (Figure 7.19). Two significant visual features are that a part of the console overlaps the main rectangle and that its lower edge is not parallel to the sides of the screen. This goes some way towards anchoring it in the diegetic space: it will be remembered that the asymmetric tendencies of the classical fiction film associate naturalism with non-parallel views (while one of Greenaway's principal devices for drawing attention to the mediation of the image was to frame it symmetrically and position it parallel to the picture plane). The technology of 1996 did not easily permit 'live' text to be presented other than parallel to the picture plane and even now there are problems of readability when texts are presented at an angle on the screen. As live text becomes

[13] Ironically, in order to achieve this effect, many of the painterly characteristics are actually water-colour painted elevations texture-mapped onto the *model*, though there is also some hand retouching of the rendered images. However, to the user this is not evident – it is the pictures which seem to be painted, not the environment. (Personal communication, Tim Warren, art director for the project, November 2000.)

both feasible and pleasant to read at other angles, then the 'naturalist' tradition of film will be likely to affect the presentation of all texts meant to be interpreted as diegetic.

Figure 7.20.
Notting Hill/JHM 1996: *The Art of Singing*
An attempt to make use of a pre-pictorial set as a site in which random access to textual information is also possible. Clicking on a topic label on a shelf triggers an animation in which a light beam travels from the shelf to the console centre-screen. Then a more or less standard text window is displayed as though it were the contents of the console screen.

In the same production, another solution is offered to the problem of combining textual and pictorial modes (Figure 7.20). Seeking to provide some of the random access to texts of an encyclopaedia, the designers encounter the problems of trying to use a certain kind of spatial realism while not imposing intolerable constraints on users – such as making them advance to each shelf and take out a book. The solution, amusing at first but soon irritating, is to connect each topic label in the library shelves to a console by a conduit through which light flows when the user clicks the label. This 'sends' information to the console which is then displayed as a more or less standard text window. Even this does not overcome the fact that the user must find the library, one specific room in a labyrinthine building, in order to get this kind of access to reference topics. Nor does it offer any of the advantages of multiple classifications and free-text searching of digital media – a book is in just one place and indexed only by simple category. It is clear that this attempt to afford multiple kinds of interaction in a single pre-pictorial space is hopelessly compromised. *Art of Singing* and similar products fail to deliver many of the functions they might offer because the spatial mode they have chosen makes it too difficult. I showed how in the case of factual television multiple media are used expressively: each media type is used appropriately and in a way which fits the other media around it,[14] but that is not the case here.

Figure 7.21.
Miller and Miller: *Myst*. Adventure Game. 1993.
All non-diegetic components are eliminated. The paper on the ground contains information which the player needs.

In Myst (**Miller and Miller 1993**) the fact that the functions to be supported are more narrow and homogenous benefits the spatial approach adopted. The environment as presented supports both idle browsing and purposive attempts to solve the 'game'

[14] This is of course not to suggest that all such decisions in television are wise ones. Postman comments at length (**Postman 1987 *passim***) on the potential damage to understanding and to mature debate caused by television's need to fill the screen with an unending series of moving pictorial images.

but need not be concerned with generalised information retrieval or reconfiguration on demand. Here also the user sees only pre-rendered views so that the advantages of optimal views – compositional and narrative interest and relevance – are again available. The user's trajectory as represented by sequences of views has a broadly naturalistic feel, with each shot having a reasonably fluid relation to its predecessor.[15] All non-diegetic elements are eliminated: if users must know something, then a note is left for them in the diegetic space (Figure 7.21).

5 Hybrid spaces

Figure 7.22.
PopTop Software 1998: *Railroad Tycoon II*. Strategy Game. The main interface.
The centre of the screen is dominated by the immersive scene while ancillary information including alternative views of the terrain is offered at the periphery.

In *Art of Singing*, the portable console was always present in the corner of the display when there were texts which could be collected from the environment. This was a step away from straightforward viewing of a pre-pictorial environment, intended to support functions which the environment itself could not. Though the object was presented as a diegetic component its spatial discontinuity from the environment tended to contradict it – it clearly is anchored to the frame of the view, not to the world. The strategy game *Railroad Tycoon* gives over a third of the display to such components, and makes similar attempts to excuse them through graphical conceits. The user never has an unencumbered view of the terrain, since it is always framed by a variety of other representations. Visually this is the most distinctive aspect of such games – their juxtaposition of many spatial schemes, styles of representation and forms of realism (Figure 7.22 and 23a). There is a map-like aerial view of the terrain, the designers having chosen to offer two optimal views on screen at once rather than making users switch between them (as noted previously, such straightforward solutions to the need to convey two kinds of information at once are taboo in film). There are configurations of buttons, compressed into the periphery to help prevent them obtruding on the main view of the terrain. Not only do the buttons have modelled shading to impart some of the familiar realism of 'objectness' but they are 'attached' to the surround using structures reminiscent of Victorian engineering in an attempt to justify their intrusion into the overall display. When a dialog-box is displayed (Figure 7.23b) the extraneous pressure gauges and dials operate and a

[15] Another aid to a sense of continuity is the use of environmental sound which is not interrupted by changes of view. This technique, noted by Persson (**1998**), is an important means by which film conceals its intermittent visual nature but is outside the scope of this thesis.

valve emits steam! To the left of the dialog it can be seen that the diegetic shadows of the trees in the landscape are indistinguishable from the extra-diegetic shadows of the dialog box itself, a literal blending of two complementary forms of realism. Even though when considered objectively this juxtaposition of realisms is disconcerting, in practice for the involved user it achieves its purpose of making a spatial hybrid more or less acceptable. However, I will suggest later that such genres may move away from this overt hybridity to a greater use of diegetic interaction.

Figure 7.23a-b. **PopTop Software 1998**: *Railroad Tycoon II*. Strategy Game.
Left, the main interface (detail). Right, a dialog box. An extraordinary amalgam of spatial styles and realisms.

Users have greater freedom to choose how they observe the pre-pictorial space than in the shot-based trajectories of the CD-ROMs using pseudofilmic space. They may look at any part of the terrain at a number of scales and may choose any of four cardinal axonometric views. The reason for limiting viewing angle and scales in this way is that every 'tile' which can appear in the terrain is pre-rendered in a series of variant forms, partly for reasons of performance but also to ensure that the user is always presented with an informative view (so once again authorial control of view is not entirely absent) but there is no expressive use of viewpoint in response to the events which unfold. Indeed such filmic responsiveness might seem impossible given that users may choose to build a railway anywhere across terrain of their own devising and that the events which may ensue are not known in advance. However, there is no reason in principle, as I suggested in relation to Kullberg's timeline, why the system itself should not select suitable variables of the view and make use of narrative spatial devices to increase the drama of the game.

Such pictorial approaches would in simple form deal only with the *permissive* aspect of film, when optimal views are allowed to the viewer. An important benefit of the game's current spatial design is that events may be unfolding out of the sight of the player and only discovered subsequently, adding to the surprise element of the game-play. Any more 'filmic' approach would need to recognise the benefits to the game-play of this invisibility, which in a filmic mode of presentation could only be dealt with by the other vital aspect of film's narrative expressivity, the deliberate denial of view. 'Intelligent' viewpoint animation is already being investigated in the

context of computer games,[16] for example to cut briefly to an aerial shot of a crashed car in racing games, but I am not aware of any automated cinematography which in addition to selecting what should be shown also decides what should not.

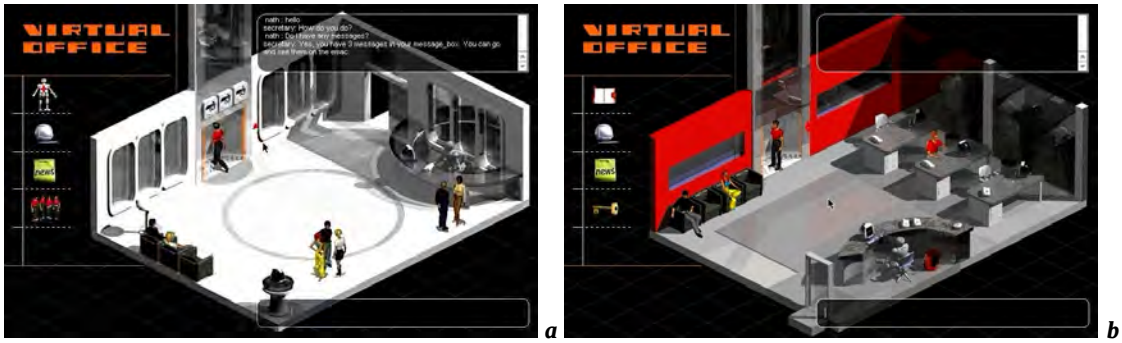
Throughout this thesis I have claimed that the selection of perspective and other projection systems, the combining of configurational and pictorial approaches, and the selection and combination of different kinds of realism in any medium, are best conceived as pragmatic, based on the objectives of the artefact, and not as matching an external referent. I remarked on the multiplicity of demands on the design process and the hybridity of the solutions which emerge. In some ways *Railroad Tycoon* is a perfect illustration: at every turn it is clear that the designers are concerned with what seems to 'work' and that the various realisms and spatial principles are selected and balanced to that end. Nevertheless, I do not wish to claim for such artefacts the maturity of expressiveness that I attributed to the fiction film. On the contrary, despite the efforts of the designers to disguise the incongruity of the components, the overall impression is that they coexist uncomfortably and the same is true of the experience of the product over time. The atmospheric, dynamic filmic sequences of the product's opening which absorb the user into a subjective view are at odds with the inert highly objective presentation which succeeds them. Perhaps Gessner's complaint that television is 'visually disorganised' *could* fairly be applied to artefacts like Tycoon. However, it is also possible that the impression of incongruity is itself a temporary phenomenon arising from the relative unfamiliarity of these spatial configurations and that in future they will come to be seen as transparent and natural. Certainly these hybrid spaces do not have the failing, so often noted in this chapter, that their purpose is unclear, and they do support well the functions which it seems they should. In that sense their expressivity is high.

Another project which is best classified as using hybrid space is a prototype office tool of primarily functional intent, though it takes into account other factors than simply executing tasks, in dealing with social as well as functional aspects of interpersonal communication in a work context. It exploits translucence as a solution to the problem of integration and moves this study towards the more fully integrated spaces of the next category.

Figure 7.24.
Chung et al 2000: Virtual Office, a project visualising a shared virtual work environment.
Translucence is used in an attempt to integrate components with different spatial modes.

[16] Personal communication from Chris Webb, Attention to Detail , UK

In *Virtual Office (***Chung et al 2000**), if a user approaches the (robot) secretary's desk (at right in Figure 7.24) a textual conversation is automatically initiated in a superimposed text chat window. Translucence binds the different modal elements closely together so that each is used appropriately to its function, but without the cost of their becoming spatially divorced. The likely importance of translucence in pictorial interactive media is discussed below. The position and visual integration of the chat window give it something of the character of a speech bubble in a cartoon (cartoons are themselves a rich area of multimodal representation deserving study as spatial artefacts), so that it seems part of the diegetic space while still being accessible to the user. Like a head-up display for a pilot it also ensures that the user may encompass diverse stimuli within a fairly narrow visual field. Axonometric views are used as in other artefacts already discussed, both for computational efficiency and to prevent loss of resolution with distance, but here they also have the effect of allowing the overlaid textual and iconic elements to be seen as more tightly integrated than if convergent perspective were used.



a ***b***

Figure 7.25a-b. **Chung et al 2000**: Virtual Office, a project visualising a shared virtual work environment. Filmic continuity devices are used: the character in red passing out of the door in ***a*** is seen entering the connected space in ***b***.

The pictorial scene functions in several ways including to indicate who is currently on line and to show what actions they are performing, and also to manage dialogue: a user may move close to another user by clicking on their avatar, and this proximity then allows text conversation to take place. Particular realisms are used, for example to represent fairly accurately the actual appearance of users, but irrelevant realisms are suppressed. Use is made of techniques derived from film to bind together separate spaces which are juxtaposed only in time and not in the display. In Figure 7.25a the user in red can be seen leaving the room by the lift, as a result of having elected to move to a different space at another level. The next scene (Figure 7.25b) shows the user emerging from the lift in that other space. No other visual information is necessary for users who are accustomed to filmic techniques.

6 Integrated spaces: combining pre-pictorial and pictorial space

In the categories of pre-pictorial and hybrid spaces discussed so far, it seemed possible to separate the model from the view to the extent that one could imagine reconstructing a coherent pre-pictorial space on the evidence of the views provided. I now analyse some spaces which at first sight also seem to offer depictions of a coherent three-dimensional space but in which it becomes apparent that the depiction in itself undermines the distinction between depiction, view and model. In this thesis

I have repeatedly cited the argument of functional mismatch as a justification for the avoidance of various realisms, and these are digital artefacts which for their different reasons deliberately abandon the realism of pre-pictorial coherence.

a

b

c

Figure 7.26a-c.
Eberle 1996: *Upholstery Weekend*.
Optimal views of the relations between 'characters' in the scene. The view in ***a*** acts as an establishing shot, creating a space using the size and geometry of the furniture items.
In ***b*** an optimal view from above allows the distance and scale of the objects to be assessed.
The chair which threatens the toy in ***b*** is seen in ***c*** in extreme close-up – a subjective shot from the point of view of the toy.

The first artefact is not interactive. This is significant and points up some important aspects of the relationship between film practice, digital space and interactivity. *Upholstery Weekend* (**Eberle 1996**) is an animated narrative. There is no attempt to abolish the awareness of representation – indeed the artefact keeps drawing attention to its mediation for ironic purposes. This is not of course a unique characteristic of digital animation – traditional animators have also made media-conscious jokes of this kind[17] – but here it helps indicate some of the ways in which mainstream film practice may be adapted and subverted in a digital space. The apparent pre-pictorial 'space' is defined only by the relationships between a small number of objects which have been dissociated from their original backgrounds. As in classical film, optimal views from widely varying viewpoints are exploited to afford the best possible view of an event or situation as well as to create visual interest and engagement. For example the shot illustrated in Figure 7.26a, in which an armchair threateningly approaches a small toy, is succeeded by that in Figure 7.26b in which the approach of the predatory armchair is more easily assessed by the viewer and in which the victim appears significantly smaller. There is a brief cut-away to an extreme close-up point-of-view shot of the front of the armchair (Figure 7.26c) before the armchair crushes the toy.

Later (Figure 7.27) defocussing (almost entirely neglected in digital media outside high-budget quasi-photographic feature films) is used to give extreme depth to the 'shot'. The impression that a second small toy is in the near foreground is a visual joke, since it turns out that this is in fact a giant toy seeking revenge.

[17] An example of such media irony in traditional animation is a character who climbs out of a hole in the ground and then picks up the black oval representing the hole and walks away with it (illustrated **Lord and Sibley 1998 p7-61**).

Figure 7.27.
Eberle 1996: *Upholstery Weekend*.
Focus is used to create distance and increase the drama of a near object.

So far this seems only a witty exercise in pre-pictorial space, but subsequently the coherence of the pre-pictorial space is subverted (Figures 7.28a-b) when a sofa is transformed in a wholly pictorial way which has nothing to do with the space which it seemed to occupy. Such 'media-savviness' will always increase the objectification of the representational artefact, here with humorous intent.

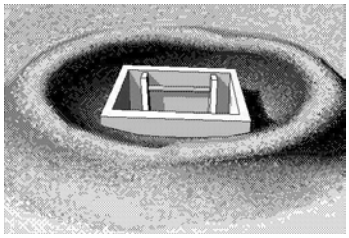
a

b

Figure 7.28a-b. **Eberle 1996**: *Upholstery Weekend*.
At times the evocation of a pre-pictorial space is abandoned in a way which forces the user to become aware of the fact of representation.

This animated narrative makes the best possible use, for its purposes, of spatial devices and their associated forms of realism, combining the particularity of photography with the digitally facilitated omission of all extraneous photographic evidence and apparently constructing, but then undermining, a pre-pictorial spatial environment. It is not simply a film delivered by digital means, because of its high level of selectivity which suppresses the accidentals of straightforward photography, and because it uses overt pictorial manipulation. However its strengths arise precisely because it is a wholly authored experience. Far from representing a solution to the difficulties of pre-pictorial spaces outlined previously, it confirms them, by underlining the expressivity of authored depictions of space in which the maker decides exactly what shall be seen, from where, with what variables of the frame, for how long, and in which every shot is designed in the known context of what precedes and follows it.

Cosmic Osmo (**Cyan Incorporated 1993**), an interactive production, also plays tricks with the user's expectations of spatial coherence. Like *Myst* (which was created subsequently by the same team) and the two Notting Hill CD-ROMs, *Cosmic Osmo* takes advantage of being limited to sequences of still images to offer optimal views of each scene. Here the 'shots' are related to one another in a way which is not intuitive


a

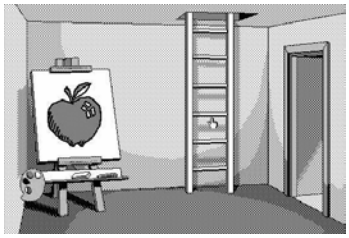

b

Figure 7.29a-b. **Cyan Incorporated 1993**: *Cosmic Osmo.*
When these 'shots' are seen consecutively, the user pauses momentarily before inferring that the ladder glimpsed in in ***a*** is that also seen in ***b***. Eventually the attempt to infer a coherent pre-pictorial space turns out to be futile.

for accustomed film-viewers. There is no attempt at 'shot-matching' so that the user is dependent on a conscious process of inference based on object-continuity. When experienced in time, rather than in space as here in Figures 7.29a-b, it takes a moment's deliberation to decide the relationship between consecutive shots. Some shot relationships are more conventional such as in Figure 7.30a-b which moves the viewpoint into the shed for a closer view.


a

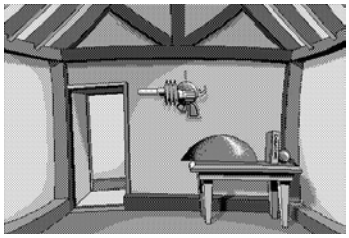

b

Figure 7.30a-b. **Cyan Incorporated 1993**: *Cosmic Osmo.*
Some shots are more conventionally related as here where the objects glimpsed inside the shed become the subject of the next scene.

Initially the user seems to explore a standard pre-pictorial space; longer exploration reveals that the space is convoluted and it becomes impossible to sustain any rational model of how the scenes are connected together. This ironical use of spatial conventions fits with the whimsical nature of the production's humour as a whole.

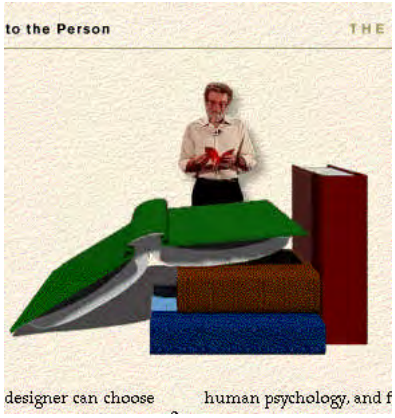


Figure 7.31.
Donald Norman 1994: *Donald A Norman – defending human attributes in the age of the machine*, CD-ROM 1994.
Integrated components dissociated from their original photographic contexts, but still partaking of some of the other realisms which photography can contribute.

Upholstery Weekend made extensive use of dissociated graphic elements extracted from their photographic or cinematographic context and positioned in a blank plane to construct a new space. Similar dissociation of photographic and quasi-photographic elements is used in 1994 in the Voyager CD-ROM of *Donald Norman* already described. In graphics like the one illustrated (Figure 7.31) an animate Don Norman is extracted from the accidentals of photography and inserted into a new space at a new scale and with shadows which belong to the environment of the

virtual page rather than to the environment in which filming took place – spatial invention which as so often selects and rejects realisms on a pragmatic basis.

Figure 7.32.
Jacques Bonnaud: film poster for *Les Enfants du Paradis* (Marcel Carné) 1944.
A compositional technique common in film posters from the 1940s. Components derived from individual production stills are collaged together in a composition which embraces multiple perspectives.
Taken from **Edwards, Gregory J,** 1985, *The International Film Poster*, Columbus Books,London, p151

Such graphical techniques are not of course confined to screen-based media. Film posters have often made use of synthetic amalgamations of photographic or quasi-photographic segments. There are a few examples of photomontage equivalent to the simple assembly class of digital collage described above, constructing almost no intra-pictorial relationships, but generally these collages take on the character of pictures in their own right, with a strong resemblance to pre-Renaissance forms of picture-making – where each component has its own optimal view (and therefore perspective geometry) largely independent of the other components, but all are nevertheless organised into a unitary composition. In the example illustrated (Figure 7.32), though the perspectives of size and geometry are not consistent with a natural scene the components are such as could be seen together in an actual pre-pictorial space, and some of the internal pictorial relationships also conform to everyday expectations (so that for example the characters look towards the woman and so do the crowds in the gallery). Below I note other forms of photomontage which take greater liberties with realism.

7 Pictorially dominated space


a

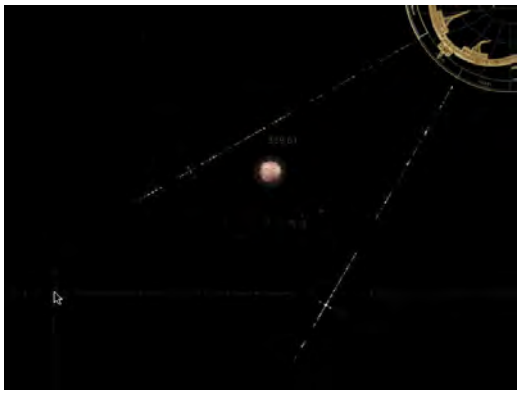

b

Figure 7.33a-b. **Holley *et al* 1998**: *Eclipse*. Objects are dissociated from their original photographic contexts and reassembled in new pictorial wholes. Objects of wide ranging scales are brought into conjunction.

Differentiated from the previous category not in kind but by the degree of pictorial arbitrariness, the final class of pictorial spaces takes further the possibilities of

recombining dissociated graphical elements into new spaces – spaces which depart further from any likely pre-pictorial model. As the integrated digital spaces described above have antecedents in montaged film-posters, so there is a tradition in other forms of photomontage of combining fragments taken from their original contexts and assembling them in a unified composition with a more or less conventional underlying pictorial structure, in some cases even evoking aspects of a naturally observed scene though those components could never appear together in the world (Figure 7.34 and numerous examples in **Ades 1986**). However, I shall suggest that interaction differentiates the case of static media from that of digital photomontage.

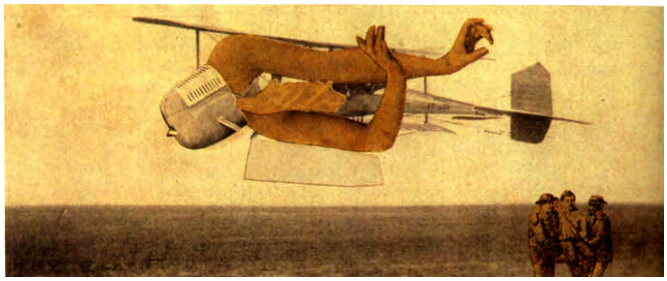


Figure 7.34. Max Ernst: *Murdering Aeroplane*, 1920, collage. From **Hughes 1980** fig38, p72.
Despite the disparate origins of their parts, many collages are contrived – often as here with ironic intent – to recall conventional compositions derived from pre-pictorial spaces.

*Eclipse (***Holley *at al* 1998***)* constructs a space from dissociated graphics, coercing disparate elements into a pictorial ensemble which, when interacted with, operates as an integrated system. Its pictorial coherence makes it unlike simple graphical assembly but also unlike a depiction of apparently pre-pictorial space. It allows an astronomical device to be brought into direct connection with the heavens which it surveys (Figure 7.33a-b), pictorially uniting elements which in pre-pictorial terms are literally light-years apart. In addition it facilitates the use of mixed modes of presentation, so that it does not seem incongruous to see spatial coordinates presented as text which alters and moves as the space changes. Graphically it gives the maker the traditional freedom of combination of photomontage but to this is added the effect of the interactivity, in which actions on one part affect the behaviour of the whole display, creating a coherence which exceeds any purely pictorial unity.

Figure 7.35.
Lac, Wölwer and Wu: Esfore-Entropy, 1998 .
Integrated components dissociated from their original photographic contexts are combined into new entities which acquire pictorial realisms of their own.

Its relation to realism is intriguing. Each graphic element has the realisms associated with photography – especially a sense of tactility arising from surface qualities and a strong sense of 'objectness' enhanced by the dissociation of elements, together with the realisms of recognition – yet the whole resembles nothing which might actually be seen in the world. Nevertheless, the sense of unity arising from the nature of the system's interactions with the user evokes the realism of coherence, a coherence

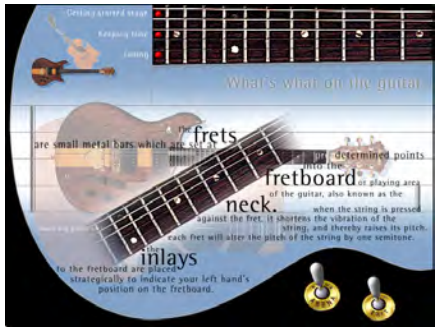
which belongs only to the pictorial and not the pre-pictorial world. *Esfore-Entropy* (**Lac, Wölwer and Wu 1998**) similarly combines dissociated graphical elements (Figure 7.35), and in this case the new ensembles move according to 'organic' patterns of behaviour, so importing an additional realism – that of animate existence. This is once again at a pictorial level, since no user will infer an actual pre-pictorial creature, even in the imaginary way that they might infer a pre-pictorial alternative life-form in a fiction film.

Taken as a whole the displays of *Eclipse* and *Esfore-Entropy* are visually unrealistic in two basic ways: they cannot be mistaken for a window on the world, and they are not automorphic to any imaginable natural scene. They *make use* of realisms but can never offer the illusion of unmediated access to the pre-pictorial world. However, interactivity does seem to make a vital difference because it reintroduces, albeit as an analogue, the ability to seek more visual information as one can in the real world (Figure 7.36). As users interact, new scenes are revealed: this tends to enable users to become at least partially unaware of the medium, even though they cannot fail to know that they are manipulating representations and not real things. Where user investigation leads to the discovery of new scenes, the naturalism of this experience seems to compensate for the evidently representational qualities of the objects and scenes encountered. There is a possible analogy here with the acceptability of editing in Film which was seen earlier to be unlike natural vision but is nevertheless perceived as in some deeper sense 'natural' and has through custom become almost invisible to the film viewer.

Figure 7.36.
Holley *et al* 1998: *Eclipse*.
A new scene is revealed through the apertures of the astronomical instrument as a result of the user's interaction.

A more prosaic use of dynamic translucence is made by *Instrumentor*, a prototype product by Art of Invention (**1995**). A display – which even in its static form makes use of translucent overlays to show more than could otherwise be accommodated in a single screen – displays additional overlays on demand to provide detailed information without loss of context (Figure 7.37a-c). This device is not without obvious difficulties: the example illustrated shows that material may become hard to decipher. However, there again seems be something of a sense of naturalness arising from the form of interaction, a feeling of direct access to information in context and on demand, which has some of the qualities of 'just seeing' even though it has no generalised resemblance to natural scenes.

a



b

c

Figure 7.37a-c.
Art of Invention 1995: *Instrumentor*.
A background graphic displays translucent overlays on demand when the user interacts with different parts of the display. Detail is displayed without loss of context.
Difficulty in reading the text and discerning the details of the graphics indicate some of the problems which may arise.

Translucence has been exploited in several of the artefacts described so far.[18] Though of some benefit in static form, it seems particularly effective when responsive to the user. I noted earlier how the use of prolonged superimposition in film was abandoned because it militated against the sense of direct access to the pre-pictorial world (while other 'non-realistic' mechanisms survived) but also pointed out its continued use as a linking device; I drew attention to the use of blended images in television to enable more than one space at a time to be combined in a single display. In digital multimedia it seems that the addition of interactivity to translucence makes a significant difference to the sense of naturalness. In existing examples the degree of translucence is at all times under authorial control, but it would be a simple matter technically for users to be allowed to control this property. A simple and direct form of interaction would be essential for this purpose. Users have been obliged to learn many new conventions and techniques in dealing with interactive digital artefacts, and this may be one which could be assimilated into a natural-seeming form of interaction. Such speculation emphasises again, as the history of film illustrated, that it is not always possible to predict which devices will come to seem natural and which will not. User-controlled translucent overlays seems to offer some promising characteristics which were not shared by the entirely authorial superimposition of early film.

A final example which also largely abandons the evocation of pre-pictorial spaces and makes extensive use of pictorial intervention is *Ceremony of Innocence* (**Mayhew 1997**). It uses the fact that it is an interactive narrative about a surreal, or perhaps magical, postcard correspondence to make visual puns on the nature of representation. Given that the narrative suggests that one of the correspondents may have invented the other, ambiguity and self-reference of representation and of space are appropriate. Since the diegetic material comprises picture postcards, everything seen on screen is

[18] Virtual Office, Eclipse and Instrumentor. It is also used, though not to any great extent, in application packages. In Macromedia Director it allows the display of ancillary information about graphical objects while not completely obscuring them.

already depictive to at least one degree (Figures 7.38a-d). The user interacts with graphic elements in ways which fluctuate between interacting with the scene depicted and interacting with the depiction itself. The bird visible in **a** is also depicted in the stamps on the card – which seems to confirm its pictorial status – yet it moves, and squawks and eats when coaxed by the pointer. The pictured wine glass in **b** breaks when hit by the pointer and the fish swims round to the other side of the card, the card turning to reveal its written side. The boundary between the diegetic and extra-diegetic is constantly invaded and the mouse pointer, normally wholly extra-diegetic and as it were the user's 'property' is often captured by, or becomes part of, the representations with which it interacts.

a *b* *c* *d*

Figure 7.38a-d. **Mayhew 1997**: Ceremony of Innocence. This CD-ROM explores the boundaries between levels of representation in which the diegetic and extra-diegetic interpenetrate.

Mayhew has said that whereas a film-maker might traditionally ensure that character development enhances the narrative while the development of the narrative enhances the understanding of character, he has attempted in *Ceremony of Innocence* to make similar use of the opportunities for the user to intervene, making a virtuous triangle of character, narrative and interaction.[19] Spatial decisions about visual representation are taken in support of this aim, so that how the representations look and how they behave are designed to be mutually appropriate. Whereas in most pictorial interactive multimedia the spatial design gives little clue to the intentions of the artefact – it is unclear why they look as they do – this is one of a small number of examples where spatiality seems to articulate content.

I do not wish to suggest that the spatiality of *Ceremony of Innocence* offers a paradigm for other kinds of interactive media – the aim of this chapter is to elucidate the

[19] Personal communication, August 1999.

different spatial forms which have emerged to serve differing requirements. *Ceremony's* objectives are relatively simple – to allow the user to progress through a narrative by means of interaction – whereas I have shown that many of the problems of devising expressive spatial organisation in interactive media arise from the conflicting demands of widely varying functions; functions which are currently visualised using the competing spatial practices of antecedent media.

I have distinguished seven classes of spatiality in pictorial interactive media and in doing so have made use of the concepts developed earlier in the thesis. This in turn has required the refinement of some of those concepts, in particular the division of spatial interventions into those concerned with the model, the view and the picture. This framework helped make sense of interactive media, clarifying the location and nature of the opportunities for spatial decisions to influence representation. However, in adding interactive media to the range of disciplines to which the framework was applied, it itself required modification to capture a different mode of spatial representation – that in which the parameters of viewing and picturing are authorially determined, but specific views and pictures are not. I discuss some implications of this below.

If the approach to spatiality proposed in this thesis is a useful one, it should not only help to explain the form of existing artefacts but also assist in showing how design may deal with new issues. These chapter conclusions therefore deal both with what may be *expected* to happen as the various genres of digital interactive media develop, and with what I suggest *should* happen, especially by indicating promising lines of inquiry in the light of this study taken as a whole. Though I have repeatedly emphasised the differences between the spatial practices of the varied visual genres which media support, differences which are profoundly tied to their objectives, I have also shown that this does not invalidate the possibility of discerning deeper correspondences between spatial practices even in widely varying genres and media.

The limitations imposed by the technologies of pictorial interactive multimedia have led to invention and a number of innovative spatial forms have been devised. Nevertheless the expressivity of the spaces of interactive media seems generally poor. While the spatialities of pictures, of film and of television seem to have a high degree of fit with their objectives, the same is not true of this newer collection of genres. Neither in terms of affect nor information are they well tuned to their objectives.

However a small number of productions seem to suggest the beginnings of new forms of expression particularly suited to pictorial interactive media which may

represent emergent genres. This question of genre is an intrinsic part of the problem. In becoming finely attuned to their objectives, the spatial forms of an axonometric engineering drawing or of a de Chirico painting have emerged as clearly identifiable genres, as have the spatiality of the fiction film and the news broadcast. The spaces of pictorial interactive multimedia are, with rare exceptions, awkward amalgams of the spatial practices of antecedent media. In general these spaces not only fail individually to serve the functions (both affective and utilitarian) of the artefact but when combined together fail to operate together in a coherent way. I have documented several instances of inappropriateness between spatial usage and design intention – or in which the intention is simply unclear.

Any discussion of the inadequacies of current spatial design for digital interactive media revealed in this chapter must take place in the light of two other insights which the thesis as a whole has made possible. The first is that there is no 'short cut' to spatial maturity – an essential component of spatial expressivity as I have shown – since this maturity lies in the relationship between representations and those who use them, not in the representation alone. The other is that the pragmatic, almost accidental, approach by which spatial innovation is achieved means that no *precise* predictions can usefully be made. What I emphasise in these concluding remarks is a general principle: the need to rethink inherited pictorial and spatial practices to suit the demands made on the artefact. This includes adapting concepts of visual realism to suit the technology and its uses.

This chapter has highlighted the influences which will determine the emergent spatiality of interactive media. Some are simple while others are more complex. The three principal ways in which digital interactive media are unlike their antecedents are: they require (at least currently) the display of everything with which the user may interact; the visual experience is only partly authored, reducing the maker's control over viewing and picturing at any moment; for the same reason, there are fewer narrative possibilities for structuring the user's experience over time. These differences affect the spatiality of the medium, or rather the different genres which the medium supports, fundamentally.

The functions of the artefact: information and affect in interactive media

Throughout the thesis I have discussed how different depth cues, realisms, and depictive devices such as illicit marks have been employed in mature media to support the objectives of the artefact. I have separated those outcomes which are principally affective from those with a stronger informational function, and shown how digital interactive media can also be understood in terms of this distinction. For example, the spatial configurations which I named *simple assembly* and *two-dimensional prepictorial space* could be used both as a practical means of giving access to multiple discrete objects, and as a means of conveying a general impression of a wealth of material. *Pseudofilmic spaces* employed a particular balance of authorial and user control to provide the user with appropriate visual information for the execution

of the task, and also by elimination of non-diegetic elements could facilitate an immersive sense of presence in a world.

In this thesis the functional demands of information and affect have sometimes been seen to be in conflict. I noted for example how the sense of co-presence and immersion is facilitated by stereopsis but that this technology offers little informational advantage so that for many purposes it has been ignored. As the design of digital interactive artefacts matures it will tend to achieve a better fit with its objectives, so that to a greater extent than now the spatial forms adopted will be those which fulfil exactly the information and affective functions required. This will lead to the abandonment of particular spatial forms where they are inappropriate, and to the increased adaptation and hybridisation of familiar kinds of spaces. This will have implications for realistic depiction, since any simple notion of realism will be undermined by functional necessity. Sometimes this has a strictly practical rationale. For example, already in most virtual environments it is possible to overcome some of the disadvantages of having to travel through the space to reach a known location by simply 'teleporting' from place to place, and I have documented several other examples in this chapter where some aspect of strong realism is rejected when it would conflict with the demands of either execution or use. In this of course it operates on principles which I have shown to be fundamental to all other forms of spatial depiction.

Crudely, one could say that affect will be a particular concern of spaces designed to seem real, present and immersive, while informational demands will dominate the design of those spaces which are overtly presentational. However this simple divergence should not be overstated. As indicated previously in the thesis, for any given artefact in use informational and affective expressivity may be closely intermingled. For example, in a three-dimensional timeline such as that of Kullberg illustrated earlier, the possibility of entering the space offers both an affective sense of immersion in time, and an informational advantage in terms of a clearer grasp of what preceded, followed or was contemporary with a particular occasion. I discuss below the relevance for interactive media of shifts in the relationship between information and affect within a single work, such as when they are sometimes in concert and sometimes in conflict in narrative.

Interaction

Though this chapter has been about the spatial design of interactive media and not about interaction as such, the importance of their mutual relationship is clear and informs most of these conclusions. Interactive media artefacts attempt to accommodate the demands of the available interaction methods, principally pointing and clicking, and this has a decisive influence on spatial design. Conversely, how an artefact responds visually to interaction influences how that system is perceived. For example, I suggested that when spatial design and interaction are highly integrated, this can turn a set of relatively unrelated visual parts into a coherent responsive system which begins to acquire a kind of naturalism of its own. It benefits from the

realism of spatial coherence even though it is not in the conventional, rather vague, sense 'realistic'. Similarly the use of superimposition as a means of providing multiple views in a single display seems to become more 'natural' when under user control, in a way that cannot be achieved in wholly authored media (I noted in Chapter 5 how it seems to be the sense of uncalled-for authorial imposition on the viewer which now makes lengthy superimpositions unacceptable in the classical fiction film). Spatial and interaction design must be conceived as complementary parts of a whole, since each affects how the other is perceived.

A sense of unmediated access to what is depicted seems to arise when an artefact requires little effort to negotiate – for example if the user need only point at some component to make it respond, rather than clicking on it; or if a display scrolls when the user approaches the perimeter rather than requiring scroll-bars or other explicit devices. This suggests the value of further exploring modes of interaction which require minimal physical intervention by the user, with a view to making selecting seem more like simply attending-to. In general it seems that 'direct manipulation' should be more direct than it is currently. Users should be able to interact directly with media content, rather than devices which in turn control it. In the short term this will lead to confusion among users, but this will not persist once new forms of interaction have become 'transparent' through familiarity. The transparency of techniques in film which seemed difficult and incongruous when they were new shows that this is possible.

Transparent and configurational approaches

What is the role of visual realism in digital interactive media? As the example of factual television showed, there is often no need to create an impression of unmediated access to the depicted subject, but in the case of film, this was exactly what was required. This is a divider of visual genres: there are those which are accepted to be primarily configurational, and those where the sense of artifice and construction is minimised (though never entirely forgotten). In digital interactive media, at one extreme lie practices in the tradition of the fiction film, designed to give in general the impression of unmediated access to reality, while at the other are the descendants of factual television and the overtly configurational approach of existing text-dominated interactive interfaces. Building on the analysis of pictures, film and television, one can predict that digital interactive media will tend to divide still more clearly between those artefacts which aspire to be taken as realistic views on a pre-pictorial world – they are intended to seem like looking at a world rather than at pictures, whether this be a fantastical environment for a game, a convincing real-world setting for an interactive fiction, an interactive documentary about a real place, or a synthetic information space – and those where overt configuration – showing, rather than the illusion of unmediated seeing – is acceptable. These are overtly representational – the user is not expected to look 'through' them but at them.

The pictorial devices of digital interactive media, while beginning already to be differentiated from those of antecedent media, will diverge from them still further.

This will also involve increasing divergence of the spatial practices of individual genres *within* digital interactive media.

When the illusion of transparent access to a pre-pictorial world is wanted, it can be predicted that those forms will tend to be accepted which, however much they actually intervene pictorially, suppress the user's awareness of that intervention by one of two methods. Either they make their intervention seem to be an aspect of the depicted model (just as film and some kinds of painting were seen to contrive diegetic excuses for authorially determined parameters of viewing and picturing), or the pictorial intervention is so chosen and used as to seem natural and transparent to the act of viewing. The analysis in preceding chapters emphasised that the sense of realism is promoted not by the presence of anything, nor by any one relationship to optical truth, but by minimising the sense of encoding or mediation. For any medium which aspires to seem realistic it will therefore be more productive to concentrate on reducing the user's awareness of mediation than on adopting any particular pictorial structure, projection system, rendering method and so forth. As repeatedly pointed out, transparency of depiction is not inherent in the artefact but arises from the interaction of the artefact and the prior knowledge embedded in the culture. In this chapter several products were described which it seemed would have benefited from a clearer decision whether to adopt an overtly presentational or a 'transparent' approach, rather than compromising awkwardly between the two. Nevertheless, since the perception of unnaturalism is aggravated by unfamiliarity, the building up of a body of shared conventions will *in itself* assist the sense of transparency.

Given the problem for those digital interactive media products which aspire to seem transparent that – in the absence of multimodal interaction – many facilities must be provided to the user through ancillary interface objects, there may be an increasing tendency to offer these operations through diegetic objects and agents, as games already tend to do. In the case of games, virtual environments, interactive films and as-yet unformed genres which aim to simulate unmediated access to visual worlds, there will be less and less toleration of extra-diegetic visual controls, such as buttons, sliders, dialog boxes and so forth.

The extreme constraints which the demands of point-and-click interaction make on the spatiality of digital interactive media will be subject to radical change when those methods are supplemented or replaced. Above all, to be able to talk to the artefact or to some agency within it will obviate the need to see the object of interest.

In relation to digital interactive media, it is probably fair to say that inherited concepts of visual realism have not yet been *sufficiently* revised. Even though digital interactive media space is heavily constrained by current interaction methods, it is clear that adoption of rigid pictorial and spatial models often unnecessarily militates against the kinds of flexibility which we associate with other, principally text-based, forms of digital media. The extent to which existing spatial forms require rethinking has been underestimated.

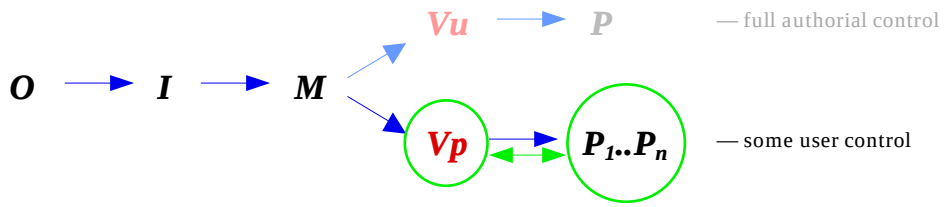
'Pictorial' aspects of configurational systems

It is not only predominantly pictorial systems which seem to require deeper exploration of the possibilities of spatial articulation. Though tangential to the main subject of this thesis, the examples of non-pictorial configuration at the beginning of this chapter suggests that such configurational systems would benefit from some of the 'sense-making' characteristics of pictures (and perhaps films) with a view to clarifying the relationships between their currently discrete parts. Rather than application software presenting the user with an uncoordinated aggregation of visual objects, some move towards constructing a coherent working 'environment' would make such software more readily comprehensible. To be beneficial these need not of course employ graphical conceits imitating the look of occupiable spaces, but visually articulate the hierarchies and interdependencies of the objects provided. As always, only relevant depth cues and pictorial devices need be employed.

Narrative versus interaction

Current quasi-cinematographic practice, as evidenced by what I have called pseudofilmic space, in which the only views seen are pre-rendered, tends to demonstrate the inexperience of current 'virtual cinematographers'. These designers would benefit from studying the details of traditional filmic practice as well as the deeper lessons which can be derived from it and which I have analysed in this thesis.

In pictures, film, television or any wholly authored genre, I have shown that the ability to choose how and when the viewer sees, including exercising outright denial of views, is a vital part of expressivity and furthers the user's engagement.



The revised framework of objective, idea, model, view and picture, which acknowledges that when dealing with prepictorial worlds it is the parameters of viewing and not necessarily the views themselves which are authorially determined, reveals that interactivity need not lead to a complete loss of opportunities for pictorial intervention. Generalised pictorial intervention is of course endemic in such systems in any case, in the sense that there must be procedures which make the model visible, but these are normally considered to be largely a technical matter. This is not surprising since it arises from that assumption of some kind of definitive 'correctness' for particular parameters of viewing and picturing which I have criticised before. Because the sense of natural vision in film in particular is so convincing, designers of digital interactive media may tend to overlook the extent to which its apparently natural qualities are dependent on authorial control.

The analysis offered here makes clear two opportunities for intervention in viewing and picturing in the interests of increased informational and affective expressivity – depictive intervention driven by the objectives of the artefact.

The first is to identify opportunities to exploit the insertion of narrative sequences into interactive software. Any moment when the user is not actually interacting can be considered such an opportunity. Such narratives can be used to clarify what is happening now and the relation to preceding and following actions, as well as to produce affective results in the viewer. This amounts to giving users views which thy did not explicitly ask for but which in the author's view will enhance the user's experience, and can clearly include spatial articulation of the model, the view or the parameters of picturing for any of the purposes which have been discussed in this thesis. Currently such narratives tend to be only carried out in the model, for example to show an avatar moving towards a place which the user has indicated.[20] But additionally, as the user moves from place to place in a virtual environment, the *view* may be altered to offer a more expressive viewpoint on the action. Or as discussed in relation to the Kullberg timeline, the *depiction* may be altered, such as throwing out of focus those items which are irrelevant to the current action. Software thus interprets the user's actions in context and adjusts viewing and picturing accordingly. Even in those systems where the viewpoint is entirely under the control of the user, there is no reason not to exploit the systematic use of picturing.

Games – prime examples of user-driven exploration of pre-pictorial spaces – show the benefits of attending to viewpoint animation, combining the expressivity of authorial choice over viewing and picturing with fluidity and freedom of exploration for the user. For example in a race game, when a character turns, the 'camera' may *follow* it, rather than rigidly tracking it. In the case where a user of a virtual environment initiates the simple action of getting into a lift, the system should extrapolate to include both entering the lift and rotating to a position facing the door. This is based on the automated sequences (such as 'dance' or 'fight') in existing virtual worlds, but unlike them is sensitive to context: the avatar does what is relevant in the circumstances. Another example of such 'context-sensitive cinematography' would be to respond to the arrival of a new character in a space in the same way that authored media habitually do – for example centring the view briefly on the new arrival before resuming a neutral position.

More sophisticated work would extend simple optimal-view based cinematography to presentational models which include an awareness of spatial narrative practice, in which the system would deduce how the content should be presented. I noted that this would include the authorial denial of view for expressive purposes, which would now be delegated to the system for decision on the author's behalf. Though of initial interest for games and interactive fictions, such techniques would eventually be

[20] Simple narratives in standard desktop interfaces include file-copy animations, the zooming of windows and the dropping down of menus. Though apparently trivial, they indicate that a scalable model of narrative which includes minimal animations at one extreme and feature films at the other (with interactive-narrative artefacts such as games somewhere in between) may prove valuable in showing how narrative interludes can be used to give greater meaning to screen-based artefacts.

applied in many interactive systems not confined to fictional genres.

Such context-sensitive cinematography would have a similar relation to realism to that of traditional cinematography, namely that while it is a deliberate pictorial intervention it is accepted as natural and even realistic because it is done on the user's behalf – it presents an answer to what Hochberg called the user's 'visual question'. Over time such motivated interventions will become transparent through acculturation.

The conclusions of this chapter have been confined to applying the observations arising from the previous chapters to a new discipline. The final chapter summarises and reviews the argument of the thesis as a whole.

8 Conclusions

1 Summary and conclusions

My experience with students and researchers had made me aware that picture-making is riddled with assumptions about what constitutes a proper picture of the world. The aim of this research has therefore been to inquire how planar representational images work. It was particularly occasioned by the suspicion that new media, especially if they are used in new ways, require, and therefore will tend to produce, new kinds of pictures. This would imply that picture-making is a pragmatic activity in which the objectives intended for the artefact are as important or perhaps more important than the relation between the picture and what it depicts, the proposition at the heart of this thesis.

The inquiry has sought a pattern of common elements across a wide range of media from static pictures, through film and television, to digital interactive media, for which there was no single body of literature. Indeed for television and interactive media there is no literature of visual representation or even of spatial practice in general. It was therefore necessary to make use of an eclectic set of literatures relevant to particular aspects of the theme, ranging through works on visual perception, theories of visual culture, film theory, art theory and history, computer graphics, and aspects of information design and human computer interaction. This ruled out the possibility of reading exhaustively the literature of any one discipline as one would in an established subject.

The development of the argument

Before embarking on the main argument of the thesis I gave a short account of two approaches to spatiality which embodied strongly contrasting approaches. One, the gestalt approach to visual perception, sought certainty in acultural factors in the apprehension of visual forms and configurations, while the other argued a high level of acculturation in the way that pictures are made and perceived. I showed that gestalt approaches, if used to inform basic design guidelines, could assist in ensuring that a spatial configuration conveyed what its designer intended, but they could not guarantee it, partly because even for quite trivial artefacts cultural knowledge had a significant effect on how the configuration was interpreted: without its assistance gestalt principles alone were inadequate.

I discussed visual culture studies partly to give a context to the debate cited later in the thesis over the extent of cultural conditioning of pictorial representation, but also for three of its insights: the raising of the question of the cultural determination of concepts of space which might otherwise seem natural and innate; the ‘invisibility’ of shared beliefs and attitudes within a culture, an idea which I adapted in this thesis to pictorial convention; and a method of dealing with the concept of intention when no conscious intending (and no particular intending individual) is implied.

Since one of the most obvious 'problems' with which pictures must deal is that they are flat – unlike the world they represent – a promising starting point seemed to be to study the depth cues used in a variety of pictures, principally paintings, from a range of periods and cultures. Using Gibson's taxonomy of depth cues, originally proposed by him partly to cast doubt on pictorial theories of vision, I was able to show that the selective use of such cues allows a varied range of results to be achieved, both in terms of information and affect. This conceptual separation – of information conveyed by a picture and changes wrought in the relationship between the picture and the user – proved useful throughout the thesis (though I emphasised that the distinction was often a difficult one to make for some aspect of any given artefact). The distinction helped clarify the answers to otherwise puzzling questions such as why some depth cues have been largely ignored and others far more generally exploited.

By using Gibson's taxonomy, I was able to show that the identification of the general concept 'perspective' with the practices of linear and size perspective had led to an impoverished view of the ways in which the depth of scenes could be represented. In particular it produced a view of depth-construction in pictures as a matter of correctness, of matching views of the world, rather than as a means to an end. I had already suggested the value of considering how each aspect of picturing might be influenced by its objectives. This led me to propose a framework for picture-making which related the objectives to several discrete stages, each involving a process of representation, which transformed the *idea* for the picture into a *model* – that which was to be depicted – which was then made visible by two further transformations, those of *viewing* and *picturing*. Though crude, this framework was to prove useful in relation to all the media under consideration, and particularly assisted in identifying the opportunities for pictorial intervention where there might otherwise have appeared to be none.

The widespread identification of geometric and size perspectives with a notion of correctness made it necessary for me to consider the concept of pictorial realism. There were a number of puzzling questions. Is there one kind of picture which is more nearly correct in its representation of scenes than any other, or perhaps is simply *right*? If one kind of picture is right, why do the great majority of pictures not conform to this form of depiction? If there is not one correct pictorial representation but many, does this mean that pictures are coded representations in the same manner as diagrams, that they achieve their effects by some means other than being 'like' what they depict? It seemed that a clear definition was required, but this proved elusive in papers devoted for example to realism in computer graphics, where visual realism was often invoked as a concept apparently requiring no definition. The implicit definitions which emerged were extremely varied, and this was to prove a problem also in relation to the alleged realism of film.

Rather than trying to use a single, unitary concept of realism, I proposed that it might be more helpful to consider a series of different realisms (I suggested a list of twelve).

This would have the benefit of making it possible to account for the sense of 'realness' created by something as simple as a drop-shadow in a computer interface, despite the fact that such a usage subscribed to almost no other aspect of real-world appearance. (If the rationale for such usages were *not* to impart some kind of sense of realness, it was difficult to see what it might be.) Like the division of depth-representation into multiple cues, the division into multiple realisms made it easier to see how each could be chosen to make a different contribution to the informational and affective outcomes of a given depiction.

Nevertheless, the question remained, whether these multiple realisms might all be brought together in one super-realist picture which was more like the scene than any other. Reviving a neglected trial of pictorial truth, I suggested that the test to answer this question should be whether a picture could create the illusion that it was not a picture but the depicted scene. However, there was always a possibility that such an illusion might be culturally determined, operating through learned codes rather than offering an acultural illusion. This required dealing with texts in the literature of visual culture which suggest, following an early suggestion by Panofsky, that many representations have equal claim to be considered correct (or that none have).

While I was able to demonstrate that there *is* a single geometry which has a superior claim to match scenes, there were a number of qualifying factors. It was obvious that true illusion could only be maintained by strictly controlling many aspects, including the content and the context – the conditions of viewing. There was considerable doubt about the possibility of establishing for certain the optical truth in relation to the non-geometrical aspects of vision, because of the complex relationship between the retinal image and the 'image' constructed from that retinal evidence by the brain. Furthermore, it seemed that while full illusion might be the only useful *theoretical* benchmark, it mattered little in practice to either picture-makers or users.

If realism-as-illusion was a limiting concept and could not account for the vast majority of figurative pictures made, this meant that the question remained unanswered of what it was that broadly realistic pictures sought to represent – and perhaps more importantly, why they did so. I proposed that if pictorial illusion, PI-realism, was not the primary aim, then perhaps what was sought was the evocation of what it was like to see the scene: visual experience or VE-realism. This would be both less and more than PI-realism: less because it could not claim any absolute truth value and might be selective in its use of pictorial devices; more because it could be used intentionally to achieve objectives that the illusory image could not, and could incorporate many known aspects of the world not available in the single optical image. Many of these could be conceived as related to time, an important aspect of vision which still pictures have generally been acknowledged to fail in imitating.

If, as I argued, VE-realism is more expressive, both in terms of information and affect, than attempts at PI-realism, it pays a price, namely that it is dependent on cultural encoding. In a way, this is just common sense: the only representation which does not

rely on shared knowledge must be the truly illusory PI-realism – any other must be harnessing prior knowledge, and probably not just of the world but of pictures too. While this might seem unfortunate, I showed that shared culture makes possible subtle visual artefacts whose spatiality supports complex interpretation. It also means that an *illusory* but still effective sense of naturalness can be created, without actually trying to imitate natural vision. This is only possible after a process of acculturation when the community of makers and users has developed shared knowledge of a body of spatial practices. This differentiates a mature from an immature medium.

In the light of the divergence between the practical and functional limitations of PI-realism and the contrasting potential of VE-realism, I proposed that we cannot say exactly to what most realistic pictures correspond, and that to seek such an answer is unnecessarily to curtail the potential expressivity of picture-making, which is better conceived as an intentional, pragmatic activity, in which decisions are made as much to achieve certain informational and affective results as matching any external referent. I argued that the spatialities of different kinds of pictures have arisen as much from *ad hoc* invention as from any coherent attempt at mimesis.

If realistic pictures use convention and other prior knowledge, how can they seem like seeing? I showed that this is where the concept of visual culture is particularly useful, emphasising the invisibility of that which is shared. It was this that could explain how a picture might seem a completely natural representation, providing unmediated visual access to a scene, within one culture, yet seem blatantly encoded in another. For the contents of a spatial representation to be effortlessly apprehended by the user, the user must be so familiar with the normal usages of the medium that there seems to be no effort of 'decoding': it is as though the user 'looks through' the spatial conventions at the subject matter itself. The sense of 'just seeing' was best characterised by the *absence* of a sense of encoding, rather than by the presence of anything.

If the literature of pictures, especially digital pictures, was full of ill-considered and conflicting definitions of realism, that of film also proved confused. In general it seemed that the kind of realism which photography contributes to the film image was assumed to hold good for the experience of film as a whole, though I showed that this is certainly not the case. When a film theorist asserted that one technique was more realistic than another, what did this mean? As with pictures, it seemed unclear what was being imitated: it was like real life, it was honest, it was like vision, it evoked the experience of being in the scene, and so forth.

I applied to film-making the *objective-idea-model-view-picture (OIMVP)* framework which I had proposed in relation to pictures. While film-making, partly because of its basis in photography, might be considered as operating serially within such a framework, so that the model is made and then filmed (which obviously in temporal terms it is), I proposed an alternative approach, in the light of the influence of the objectives on every aspect of the depiction, in which the *whole process* could be

considered essentially pictorial, since the intention at every stage was to achieve a series of images from which the user could construct a meaningful narrative. In that sense every part of the process was a contribution to the (possibly false) pictorial evidence provided to the viewer.

Just as for pictures, the question of *seeming real* arose. If, as I had shown, film was fundamentally contrived in the service of its narrative objectives, how could it seem like seeing? By evaluating techniques which had at one time been popular with film-makers but had eventually been rejected for the classical fiction film, I was able to discern a pattern. Just as one would have expected from the analysis of pictures, the techniques eliminated were those which drew attention to the mediation of the scene. I used the contrast with a style of film-making whose objectives were quite different – and also that of some kinds of factual television – to point up the difference between those representational techniques chosen because of their relatively low visibility, their ability to become transparent to the content and not to obtrude a sense of mediation, and those which are overtly presentational.

If film is driven more by its objectives than by mimesis, what are they? I showed that the spatiality of film is essentially subservient to narrative. However, not only was it clear that film serves narrative, but also that narrative serves film, in the sense that it is another important means by which its unlikeness to natural vision is obscured. Because film, by use of what I called the *optimal view*, answers the question which the viewer has been prompted to ask, it acquires a sense of naturalism which is mistaken for being like seeing even while in fact it is a highly authored form of showing. However, as for all depiction, there is no single, simple answer. Just as it is not the remit of most pictures to be wholly transparent, since much of the affect of pictures comes from the viewer's dual awareness of them as substance and as 'window', certain kinds of awareness of representation in film are also important. Authorial denial of view and other obtrusions into the sense of 'just seeing' emphasised still further how the spatiality of film is bound up with its narrative objectives.

The analysis of the spatiality of television, a subject without an existing literature, allowed two main points to be made. One was that despite lacking the sense of unity and naturalness of fiction film, the space of factual television is equally well attuned to its objectives. The other was that even within the genre of factual television broadcasting, there is variation of spatial practice, depending – of course – on its objectives. I related these differences to the *OIMVP* framework. Whereas some programme-making makes uninhibited use of overt configuration within the pictorial display (and in so doing approaches the margins of the subject of this thesis) I showed how other programmes make pictorially determined interventions which either belong, or are disguised as belonging, to the model. In so doing they tend towards the unitary image of an apparently preexisting model familiar from the classical film.

Finally I considered the uses of spatiality in digital interactive media, applying to it the analysis of depiction developed in the preceding chapters. This allowed me both to make sense of the variety of existing practice and to discern opportunities for future development.

Running through the preceding part of the thesis had been an argument that the spatial practices of pictures, film and television could be considered mature. By contrast, digital interactive media (at least when attempting to make use of depiction) seems immature, placing undue reliance on inherited practices from antecedent media, and suffering from the lack of shared knowledge common to makers and users which would make its pictorial usages seem transparent.

I have demonstrated that picture-making in all media is best conceived as a pragmatic activity in which the objectives intended for the artefact are as important as, often more important than, the relation between the picture and what it depicts. This is of course itself a pragmatic argument: when I say 'best conceived' I mean that it serves two useful purposes. It can explain the extraordinary variety of depictions which are accounted 'realistic'. It is also the most useful model for the designer of pictorial representations, since it clarifies what purposes pictures serve and how they serve them. The practical implications for the design of digital interactive media have been drawn out in the previous chapter.

Several subsidiary principles have emerged during the course of the main argument, which I summarise here. They all in some way have implications for design. While I feel that I have 'proved' the main argument of the thesis, some of these observations are advanced more tentatively.

2 Implications for design

Designers should not fear innovation on the grounds that users will not understand new spatial forms, interaction techniques, and so forth since, as the example of film showed, spatial practices in immature genres are often difficult to understand. Media genres such as film would not exist in their present highly expressive form if early practitioners had limited themselves to using techniques already fully understood by their viewers.

The transition to maturity is more likely to be achieved by *ad hoc* innovations designed to solve particular needs, than by adherence to rigid systems and standards. Inventions made to overcome technical impediments are not necessarily abandoned when those impediments are taken away.

Pragmatism

Spatial design should be approached pragmatically. The key criterion is what *works* – that is, what works for a given set of objectives and users and in a given context (including cultural context) – rather than the matching of some external referent such as a scene. This has obvious implications for realism. In some cases the main objective

of a representation will be to convince the user that what they see is like reality, but that does not mean that in any simple sense it need be.

Spatial design is pragmatic in another sense too: it is best conceived in terms of *what users can do with it* and *what it does to users*. The key consideration is the range of objectives which the artefact must fulfil.

Realism

Realism is not the uncontentious unitary concept it is often presented as. There are two problems with simplistic concepts of realism: we cannot easily say what in the world we are trying to capture when we make spatial representations, nor that some kinds of 'capture' are definitively correct. It is better understood as multiple realisms which can be selected according to suitability to the objectives of the design.

If there is not a simple, easily defined realism, then the realisms used should be defined in terms of their reception by the user, as much as in terms of their relation to an original. As with depth cues, more realisms are not necessarily better: expressivity may be increased by their selective use.

One of the most basic achievements of a graphical representation is to suggest depth in a planar image. However, the view of 'progress' since the Renaissance is unhelpful when it prioritises geometry (linear and size perspectives) over other important depth cues. Even for images which aim to produce a sense of transparent viewing of a pre-pictorial world, not all depth cues will be either necessary or desirable, since different depth cues afford different kinds of looking at, and interacting with, images.

Though there is one projection system which offers a stimulus to the retinae which is closest to that from natural scenes, this is often irrelevant to the expressivity of an image.

Expressivity

Picture-making is a process of transformation in which the task is to *make* something from the visual material, selecting which aspects of scenes in the world are to be represented (in some way) in the artefact. Complete representation is impossible and unnecessary. Omission is as important as inclusion in expressivity. For example, all film-making relies on omission – framing images and selecting shots – for its expressive power. In informational terms, the suppression of detail in a picture, or the adoption of a non-optical projection system, may make an image more informative: this also is expressivity. However, representations rely on other strategies as well: marks and other interventions are inserted in the artefact which are not visible in the scene ('illicit marks'); most representations do not simply evoke the raw optical impression of a scene ('PI-realism') but some kind of experience of seeing ('VE-realism'). Even if a picture-maker *were* trying to reproduce the raw optical impression, there is a limit to the certainty with which it can be said what that raw impression is.

In specifying and designing pictures there is no place for 'purist' approaches. The picture-maker using paint, photography, computer graphics, still, moving or interactive images should consider any picturing as a process of *creating* an experience for the user.

Distinguishing information and affect

An important difference among objectives is between the affective and the informational. The objectives of any artefact should be considered as including what the maker wants to do to, or for, the user and therefore what relation the user should have to the artefact. Crudely, informational expressivity is a measure of how much the user knows more about the depicted scene, while affective expressivity reflects changes to the relation between the user and the scene.

Some forms of spatial representation which add little in the way of information may nevertheless make a fundamental difference to affect (for example binocular imagery). Often however a spatial attribute will contribute to both affective and informational aspects (for example a close-up which allows a face to be studied in detail also causes it to impinge on the user's space).

By understanding what each of the spatial characteristics is capable of contributing in terms of information or affect to the reception of a given artefact, greater expressivity can be achieved.

Maturity

In a mature medium, limitations are exploited as strengths. For example, the way scenes are clipped by the frame in traditional pictures and in established screen-based media has been extensively exploited as part of their expressivity. In a mature medium there is a reasonable number of forms available with which to express any meaning which may arise and these forms are sufficiently pliable to be modified and combined in order to suit the objectives. The form is a good fit with the objectives: when the user experiences an artefact in a mature medium, the objectives of the maker are clearly inferred (probably unconsciously) from the design. The role of the user is also crucial because a mature medium is one in which a community of shared knowledge has arisen between the makers and the users of the medium, allowing its conventions to be largely unnoticed.

In an immature medium, techniques are noticed, and this act of noticing gets in the way of any direct, natural sense of 'just seeing'. Even in traditional media, a newly introduced technique may remain problematic and be generally rejected, or it may become assimilated into the range of familiar and therefore transparent practices. However, too-noticeable representational practices may still find a place in more overtly configurational genres. For new spatial forms to become part of an expressive 'language', use must reveal a close relationship between the form, the maker's intentions and the user's needs.

Maturity does not lie in media or technology. It is a measure of how well developed the relationship is between a technology and its users. There is therefore no instant solution: it inevitably takes time for the genres allowed by a technology to reach maturity.

Genre

Spatial characteristics vary according to genre. The relation between genre and characteristic spatiality is a symbiotic one in which each helps to define the other.

For any genre over time, only those spatial devices which are expressive will survive.

Stephen Boyd Davis
1 June 2002

9 Bibliography

Texts

Ades, Dawn, 1986, Revised and enlarged from edition of 1976, *Photomontage*, Thames and Hudson, London

Allen, Richard, 1995, *Projecting Illusion: film spectatorship and the impression of reality*, Cambridge University Press, Cambridge, UK

Alpers, Svetlana, 1983, *The Art of Describing – Dutch Art in the Seventeenth Century*, University of Chicago Press / John Murray, London

Andersen, PD, 1990, *A Theory of Computer Semiotics: semiotic approaches to construction and assessment of computer systems*, Cambridge University Press, Cambridge, UK

Armes, Roy, 1994, *Action and Image: dramatic structure in cinema*, Manchester University Press, Manchester, UK

Arnheim, Rudolf, 1956, *Art and Visual Perception: a psychology of the creative eye*, Faber and Faber, London

Arnheim, Rudolf, 1993, 'Sketching and the psychology of design', *Design Issues*, Spring 1993, Vol. IX, No. 2

Ascott, Roy, 1994, www, 'The Architecture of Cyberception', *Proceedings of ISEA'94*, 5th International Symposium on Electronic Art, Helsinki, Finland, published at http://caiiamind.nsad.newport.ac.uk/cyberception.html (1994)

Bablet, D, 1966 (1962), *Edward Gordon Craig*, Published in French 1962, Translation Heinemann 1966, Heinemann, London

Bailey, Fiona and Moar, Magnus, 2000, 'Children's Creation of Shared 3D Worlds', Proceedings of Conference, *Digital Content Creation*, Bradford, April 2000, also book: Springer, Berlin due for publication 2001.

Baker, Robin, 1993, *Designing the Future: the computer transformation of reality*, Thames and Hudson, London

Bann, Stephen, 1987, 'Art', in Cohn-Sherbok, Dan and Irwin, Michael (eds) *Exploring Reality*, Allen and Unwin, Boston, 83-108

Barlow, Horace, 1990, 'What does the brain see? How does it understand?', in Barlow, Horace; Blakemore, Colin and Weston-Smith, Miranda (eds) *Images and Understanding*, Cambridge University Press, Cambridge, p5-25

Barthes, Roland, 1977, *Image – Music – Text* (translated from the French by Stephen Heath 1977), Fontana, London

Barthes, Roland, 1973 (1957), *Mythologies*, translated from the French (Mythologies, Éditions du Seuil 1957) by Annette Lavers 1972, Paladin Grafton, London

Baxandall, Michael, 1985, *Patterns of intention – on the historical explanation of pictures*, Yale University, New Haven and London

Baxandall, Michael, 1995, *Shadows and Enlightenment*, Yale University Press, New Haven and London

Bazin, André, 1967, *What is Cinema?* Volume 1, trans. Hugh Gray. Originally published in Editions du Cerf, Paris: Qu'est-ce que le Cinéma? in 4 vols 1958-1965, University of California Press, Berkeley

Becker, Alton L, 1995, *Beyond translation: essays toward a modern philology*, University of Michigan Press, Ann Arbor

Bertin, Jacques, 1973, 2nd edition, *Sémiologie Graphique: les diagrammes, les résaux, les cartes*, Mouton / Gauthier-Villars, Paris

Bertin, Jacques, 1983, *Semiology of graphics: diagrams, networks, maps* (2nd Edition of Sémiologie graphique, 1973, translated by William J Berg), University of Wisconsin Press, Madison

Birrell, Simon, 1999, 'Personal Reflections on the Development of Cyberspace', in Jacobson, Robert (ed.) *Information Design*, MIT Press, Cambridge MA, p327-338

Bordwell, David, 1985, 'Space in the Classical Film', in Bordwell, David; Staiger, Janet and Thompson, Kristin 1985 *The Classical Holywood Cinema*, Routledge, London, p50-59

Bordwell, David, 1976, 'Citizen Kane', in Gottesman, Ronald (ed.) *Focus on Orson Welles*, Prentice-Hall, Englewood Cliffs, NJ, 103-125

Bordwell, David; Staiger, Janet and Thompson, Kristin, 1985, *The Classical Holywood* Cinema, Routledge, London

Boyd Davis, Stephen, 2000, *Media Space: the uses of spatiality*; a report commissioned by BTexaCT (Research Laboratories of BT), December 2000, 250 pages

Boyd Davis, Stephen and Athoussaki, Helena, 1999, 'VRML: a Designer's view', in Vince, John and Earnshaw, Rae *Virtual Worlds on the Internet* Proceedings of Virtual Environments Conference, Bradford, 15-16 April 1997, IEEE Computer Society, p35-51

Boyd Davis, Stephen; Lansdown, John and Huxor, Avon, 1997, *The Design of Virtual Environments*, Report for the Support Initiative for Multimedia Applications of the JISC New Technologies Initiative, SIMA Report No 27, July 1997, ISSN 1356-5370

Brownlow, Kevin, 1968, *The Parade's Gone By*, University of California Press, Berkeley

Brownlow, Kevin, 1983, *Napoleon – Abel Gance's classic film*, Jonathan Cape, London

Bruce, Vicki; Green, Patrick R and Georgeson, Mark A, 1996, 3rd edn, *Visual Perception: physiology, psychology and ecology*, 1996, Psychology Press (Taylor and Francis), Hove, East Sussex, UK

Bryson, Norman, 1990, *Looking at the Overlooked: four essays on still-life painting*, Reaktion Books, London

Burton, Edward, 1995, 'Thoughtful Drawings: a computational model of the cognitive nature of children's drawing', *Proceedings of Eurographics '95*; Maastricht, NL, August 28 - September 1, 1995, C159-C170

Callaghan, Barry, 1973, *The Thames and Hudson Manual of Film-making*, Thames and Hudson, London

Callow, Simon, 1995, *Orson Welles: the road to Xanadu*, Jonathan Cape

Card, Stuart K; Mackinlay, Jock D and Shneiderman, Ben, 1999, eds., *Readings in Information Visualization: using vision to think*, Morgan Kaufmann Publishers, San Francisco

Card, Stuart K; Pirolli, Peter and Mackinlay, Jock D, 1994, 'The Cost-of-Knowledge Characteristic Function: display evaluation for direct-walk information visualisations', *Proceedings of CHI'94*, ACM Conference on Human Factors in Computing Systems, Boston, ACM, New York, p238-244

Carringer, Robert L, 1996, revised edition, *The Making of Citizen Kane*, University of California Press, Berkeley

Carroll, Noël, 1996, *Theorising the Moving Image*, CUP, Cambridge, UK

Cassell, J; Bickmore, T; Billinghurst, M; Campbell, L; Chang, K; Vilhjálmsson, H and Yan, H, 1999, 'Embodiment in conversational interfaces: Rea', *Proceeding of the CHI 99 conference on Human factors in computing systems: the CHI is the limit* May 15 - 20, 1999, Pittsburgh, PA USA, ACM, New York, p520-527

Chapman, Paul; Stevens, Peter; Wills, Derek and Brookes, Graham, 1998, 'Seabed visualisation', *IEEE Proceedings of the conference on Visualization '98*, 1998, ACM, New York, pp479-481 and 572

Clark, Kenneth, 1969, *Civilisation: a personal view*, British Broadcasting Corporation and John Murray

Cole, Alison, 1992, *Eyewitness Art: Perspective*, Dorling Kindersley / National Gallery Publications, London

Coleridge, Samuel Taylor, 1817, *Biographia Literaria*, 2 vol. Rest Fenner, London

Crary, Jonathan, 1990, *Techniques of the Observer: on vision and modernity in the nineteenth century*, MIT Press,

Crook, Geoffrey, 1986, *The Changing Image: television graphics from caption card to computer*, Built by Robots Press, London

Currie, Gregory, 1995, *Image and Mind: film, philosophy and cognitive science*, Cambridge University Press, Cambridge UK

Deregowski, Jan B, 1984, *Distortion in Art: the eye and the mind*, Routledge and Kegan Paul, London

Diefenbach, Paul J and Badler, Norman I, 1997, 'Multi-Pass Pipeline Rendering: Realism For Dynamic Environments', *Proceedings, 1997 Symposium on Interactive 3D Graphics*, Providence RI, USA, ACM, New York, pp59-70

Dorling, Daniel and Fairbairn, David, 1997, *Mapping: ways of representing the world*, Addison Wesley Longman, Harlow, England

Drake, Stillman, 1986, 'Literacy and Scientific Notations', Wrolstad, Merald E and Fisher, Dennis F *Towards a New Understanding of Literacy*, Praeger, New York, p135-155

Druks, Eldad, 1995, 'De Chirico City', *Computers and the History of Art*, 1995, Vol 5, No 2, Harwood Academic Publishers GmbH, p107-119

Dubery, Fred and Willats, John, 1983, *Perspective and other drawing systems*, The Herbert Press, London

Dunbar, Robin, 1996, *Grooming, Gossip and the Evolution of Language*, Faber and Faber, London

Dunning, William V, 1991, *Changing Images of Pictorial Space: a history of spatial illusion in painting*, Syracuse University Press, New York

Dykes, JA; Moore, KE and Fairbairn, D, 1999, 'From Chernoff to lmhof and Beyond: VRML and Cartography', *Proceedings of VRML 99*, Paderborn, Germany 1999, ACM, New York, pp99-103

Edgerton, Samuel Y, Jr, 1976, *The Renaissance Rediscovery of Linear Perspective*, Icon / Harper and Row, New York

Eisenstein, Sergei, 1977 (1949), *Film Form: essays in film theory*, Edited and translated by Jay Leyda, Harcourt Brace, Orlando, Florida

Elkins, James, 1994, *The Poetics of Perspective*, Cornell University Press, Ithaca NY

Ellis, John, 1992, 2nd Edition, *Visible Fictions: cinema, television, video*, Routledge, London and New York

Ellis, Les, 2000, 'Battle of the babes', *Computer and Video Games*, May 2000, No.222, p36-7

Elsaesser, Thomas (with Adam Barker), 1990, eds., *Early Cinema – Space, Frame, Narrative*, BFI Publishing, London

Englander, A Arthur and Petzold, Paul, 1976, *Filming for Television*, Focal Press, London

Evans, Jessica and Hall, Stuart, 1999, 'What is Visual Culture?', in *Visual Culture: the reader*, Sage Publications / Open University, London

Evinger, Craig; Manning, Karen A; Pellegrini, John J; Basso, Michele A; Powers, Alice S and Sibony, Patrick A, 1994, 'Not looking while leaping: the linkage of blinking and saccadic gaze shifts', *Experimental Brain Research* 1994, Vol.100 No.2, p337-44

Fischetti, Mark, 2000, 'The Future of Digital Entertainment', *Scientific American*, November 2000, Vol.283, No.5, p31-33

Foley, James D; van Dam, Andries; Feiner, Steven K and Hughes, John F, 1995, *Computer Graphics: principles and practice, Second Edition in C*, Addison Wesley, Reading, MA

Foucault, Michel, 1974 (1966), *The Order of Things* (translated 1970 from Les Mots et les Choses published by Gallimard 1966), Routledge, London

Frizot, Michel, 1998, (ed.), *A New History of Photography*, Könemann, Köln

Furnham, David, 1999, *Documentary Practice*, unpublished PhD thesis at Middlesex University, UK, February 1999

Garland, Ken, 1994, *Mr Beck's Underground Map*, Capital Transport Publishing, Middlesex, UK

Gautrand, Jean-Claude, 1998b, 'Stereoscopy', in Frizot, Michel, 1998 (ed.) *A New History of Photography*, Könemann, Köln, p178

Gautrand, Jean-Claude, 1998a, 'Photography on the Spur of the Moment: instant impressions', in Frizot, Michel, 1998 (ed.) *A New History of Photography*, Könemann, Köln, p233-241

Gaver, William, 1992, 'The Affordances of Media Spaces for Collaboration', *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, November 1992, ACM, New York, p17-24

Gell, Alfred, 1998, *Art and Agency*, Oxford University Press, Oxford

Gessner, Robert, 1968, *The Moving Image: a guide to cinematic literacy*, Cassell, London

Gibson, James J, 1950, *The Perception of the Visual World*, Houghton Mifflin, Boston MA

Gibson, James J, 1954 'A theory of pictorial perception', *Audio Visual Communication Review*, 2(1), p3-23.

Gibson, James J, 1979, *The Ecological Approach to Visual Perception*, Houghton Mifflin, New York

Gibson, William, 1986, *Neuromancer*, Grafton Books, London

Giejgo, Marja, 2001, www, 'Val Geilgud and the BBC', website of *Independent Radio Drama Productions Ltd*, http://www.irdp.co.uk/GIELGUD/valbbc14.htm (1 September 2001)

Gombrich, Ernst H, 1973, 'Illusion and Art', in Gregory, Richard L and Gombrich, Ernst H, *Illusion in Nature and Art*, Duckworth, London, p193-243

Gombrich, Ernst H, 1980, 'Standards of Truth: the arrested image and the moving eye', in Mitchell, WJT (ed.) *The Language of Images*, University of Chicago Press, Chicago, p181-217

Gombrich, Ernst H, 1977, 5th edn., *Art and Illusion: a study in the psychology of pictorial representation*, Phaidon, London

Goodman, Nelson, 1969, *Languages of Art: an approach to a theory of symbols*, Oxford University Press, London

Goody, Jack, 1987, *The Interface between the Written and the Oral*, Cambridge University Press, Cambridge

Granieri, John P; Crabtree, Jonathan and Badler, Norman I, 1995, 'Production and playback of human figure motion for visual simulation', *ACM Transactions on Modelling, Computation and Simulation*, July 1995, Vol 5, No 3, ACM, New York, p222-241

Grasso, Michael A; Ebert, David S and Finin, Timothy W, 1998, 'The integrality of speech in multimodal interfaces', *ACM Transactions on Computer-Human Interaction*, August 1995, Vol.5, No.4, ACM, New York, p303-325

Greenberg, Donald P., 1999, 'A Framework for Realistic Image Synthesis', *Communications of the ACM*, August 1999, Vol. 42, No. 8, ACM, New York, p44-53

Gregory, Richard L, 1970, *The Intelligent Eye*, Weidenfeld and Nicolson, London

Gregory, Richard L, 1977, 3rd Edition, *Eye and Brain*, Weidenfeld and Nicolson, London

Gregory, Richard L, 1987, Perception as Hypotheses, in Gregory, RL (ed.) *The Oxford Companion to the Mind*, Oxford University Press, Oxford, p608-611

Gregory, Richard L, 1998, 5th Edition, *Eye and Brain*, Oxford University Press, Oxford

Hagen, Margaret A, 1980, 'Generative Theory: a perceptual theory of pictorial representation', *The Perception of Pictures: Vol II*, Academic Press, New York, p3-46

Hall, Edward T, 1966, *The Hidden Dimension: man's use of space in public and private*, The Bodley Head, London

Hall, Edward T, 1959, *The Silent Language*, Doubleday, Garden City

Hanson, Anne Coffin, 1979, 2nd edn., *Manet and the Modern Tradition*, Yale University Press, New Haven

Harrington, John, 1973, *The Rhetoric of Film*, Holt, Rinehart and Winston, New York

Harrison, Steve and Dourish, Paul, 1996, 'Re-Place-ing Space: The roles of place and space in collaborative systems', *Proceedings of CSCW'96*, ACM, New York, p67-76

Hart, Vaughan and Day, Alan, 1995, 'A Computer Model of the Theatre of Sebstiano Serlio, 1545', *Computers and the History of Art*, 1995, Vol 5, No 1, Harwood Academic Publishers, p41-52

Hart, Vaughan and Robson, Joe, 1999, 'Carlo Crivelli's Annunciation (1486) A Computer Investigation into Renaissance Painted Perspective', *Computers and the History of Art*, 1999, Vol 8, No 2, Harwood Academic Publishers, p55-70

Harvey, David, 1996, *Justice, Nature and the Geography of Difference*, Blackwell, Malden MA

He, Li-wei; Cohen, Michael F. and Salesin, David H., 1996, 'The virtual Cinematographer: a paradigm for automatic real-time camera control and directing', *Proceedings of the 23rd annual conference on computer graphics* August 4 - 9, 1996, New Orleans, LA USA, International Conference on Computer Graphics and Interactive Techniques, p217-224

Healey, Christopher G, 1999, www, Preattentive Processing, http://www.csc.ncsu.edu/research/areas/hci.html (12 June 2000)

Heidrich, Wolfgang; Westermann, Rüdiger; Seidel, Hans-Peter and Ertl, Thomas, 1999, 'Applications of pixel textures in visualization and realistic image synthesis', *Proceedings of the 1999 symposium on Interactive 3D graphics*, 1999, ACM, New York, p127-134

Hochberg, Julian, 1980, 'Pictorial Functions and Perceptual Structures', *The Perception of Pictures: Vol II*, Academic Press, New York, p47-94

Hochberg, Julian, 1987, 'Perception of Motion Pictures', in Gregory, RL (ed.) *The Oxford Companion to the Mind*, Oxford University Press, Oxford, p604-8

Hodges, Andrew, 1983, *Alan Turing: the Enigma of Intelligence*, Burnett Books/Hutchinson 1983, republished Unwin Paperbacks 1985, London

Hoffman, Donald D, 1998, *Visual intelligence: how we create what we see*, WW Norton and Company, New York and London

Hogben, Lancelot, 1949, *From Cave Painting to Comic Strip: a kaleidoscope of human communication*, Max Parrish, London

Hopgood, FRA, 1993, 'Use of time and space in multimedia systems', *Multimedia in Higher Education: portability and networking* AGOCG Technical Report No. 24 December 1993, 7 pages [no page numbers]

Houlgate, Stephen, 1993, 'Vision, Reflection, and Openness', in Levin (ed.) *Modernity and the hegemony of vision*, University of California Press, Berkeley and Los Angeles, CA, 87-123

Hughes, Robert, 1980, *The Shock of the New: art and the century of change*, British Broadcasting Corporation, London

Ivins, William M, 1975 (1938), *On the Rationalization of Sight* (originally published by Metropolitan Museum of Art 1938), Da Capo Press, New York

Ivins, William M, Jnr, 1953, 'The Blocked Road to Pictorial Communication', in Harrison, Charles and Orton, Fred (eds.) *Modernism, Criticism, Realism* 1984, Harper & Row, London, UK

Jackson, Peter, 1993, 'Towards a cultural politics of consumption', in Jon Bird, Barry Curtis, Tim Putnam, George Robertson and Lisa Tickner (eds.) 1993 *Mapping the Futures: local cultures, global change* (based on a conference 'Futures' November 1990 Tate Gallery, London) Routledge London, pp207-228

Jenks, Chris, 1995, 'The Centrality of the Eye in Western Culture: an introduction', in Jenks, Chris (ed.) 1995 *Visual Culture*, Routledge, London, p1-25

Jones, Huw, 2001 (in press), *Computer Graphics through Key Mathematics*, Springer, London

Jones, Rebecca K and Hagen, Margaret A, 1980, 'A Perspective on Cross-Cultural Picture Perception', *The Perception of Pictures: Vol II*, Academic Press, New York, p193-226

Kandogan, Eser and Shneiderman, Ben, 1997, 'Elastic Windows: Evaluation of Multi-Window Operations', *Proceedings of ACM SIGCHI 97 Conference on Human Factors in Computing Systems* (March 1997), ACM, New York, p250-257

Katz, Ephraim, 1994, 2nd edition, ed., *The Macmillan International Film Encyclopedia*, Pan Macmillan, London

Klee, Paul, 1968 (1925), *Pedagogical Sketchbook* (translated by Sybil Moholy Nagy from the Pädagogisches Skizzenbuch 1925), Faber and Faber, London

Kress, Gunther and van Leeuwen, Theo, 1996, *Reading Images: the Grammar of Visual Design*, Routledge, London

Kuhn, Thomas, 1970, 2nd edn (1st edn 1962), *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago

Kullberg, Robin L., 1995, www, *Dynamic Timelines: Visualizing Historical Information in Three Dimensions*, Thesis, Master of Science in Media Arts and Sciences, Massachusetts Institute of Technology, http://robin.www.media.mit.edu/people/robin/thesis/ (at April 1999),

Kurlander, David; Skelly, Tim and Salesin, David, 1996, 'Comic Chat', *Proceedings of the 23rd Annual Conference on Computer Graphics*, 1996, ACM, New York, p225-236

Lakoff, George and Johnson, Mark, 1980, *Metaphors We Live By*, University of Chicago Press, Chicago

Lalvani, Suren, 1996, *Photography, vision and the production of modern bodies*, State University of New York Press, New York

Lannoch, Helga and Lannoch, Hans-Jurgen, 1989, 'Towards a Semantic Notion of Space', *Design Issues*, Spring 1989, Vol.5, No.2

Lansdown, John, 1987, *Computer Graphics*, Hodder and Stoughton, Sevenoaks, Kent, UK

Lansdown, John and Schofield, Simon, 1995, 'Expressive rendering: a review of nonphotorealistic techniques', *IEEE Computer Graphics and Applications*, Vol.15, No.3, p29-37

Levin, David Michael, 1993, ed., *Modernity and the hegemony of vision*, University of California Press, Berkeley and Los Angeles, CA

Lie, Håkon Wium and Bos, Bert, 1997, *Cascading Style Sheets: designing for the Web*, Addison Wesley Longman

Lord, Peter and Sibley, Brian, 1998, *Cracking Animation – the Aardman book of animation*, Thames and Hudson, London

Lothe, Jakob, 2000, *Narrative in Fiction and Film*, Oxford University Press, Oxford

Macdonald, Gus, 1979, *Camera – a Victorian eyewitness*, Batsford, London

Mackinlay, Jock D, 1986, 'Automating the Design of Graphical Presentations of Relational Information', *ACM Transactions on Graphics*, Vol 5, No 2, ACM, p110-141

Markus, Thomas A, 1993, *Buildings and power – freedom and control in the origin of modern building types*, Routledge, London

Marr, D and Nishihara, H., 1978, 'Representation and Recognition of the Spatial Organisation of Three-dimensional Shapes', *Proceedings of the Royal Society of London*, Vol. B. 200, The Royal Society, London, p269-294

Marr, David, 1980, 'Visual Information Processing: the structure and creation of visual representations', in Longuet-Higgins, HC and Sutherland NS (eds.) *The Psychology of Vision* (a Royal Society discussion organized by HC Longuet-Higgins, FRS, and NS Sutherland, held on 7 and 8 March 1979), The Royal Society, London, p199-218

Marr, David, 1982, *Vision*, WH Freeman and Company, New York

Marshall, Catherine C and Shipman, Frank M, III, 1995, 'Spatial Hypertext: Designing for Change', *Communications of the ACM*, August 1995, Vol. 38, No. 8, ACM, New York, p88-97

Marvin, Carolyn, 1988, *When Old Technologies Were New – thinking about communications in the late nineteenth century*, Oxford University Press, Oxford

Massey, Doreen, 1993, 'Power-geometry and a progressive sense of place', in Jon Bird, Barry Curtis, Tim Putnam, George Robertson and Lisa Tickner (eds.) 1993 *Mapping the Futures: local cultures, global change* (based on a conference 'Futures' November 1990 Tate Gallery, London) Routledge London, pp59-69

McBride, Joseph, 1972, *Orson Welles*, Secker and Warburg / BFI, London

Melia, Paul and Woods, Alan, [1998] (undated), *Peter Greenaway: artworks 63-98*, Manchester University Press / Cornerhouse gallery, Manchester UK

Merritt, Douglas, 1993, *Graphic Design in Television*, Focal Press / Butterworth-Heinemann, Oxford

Merritt, Douglas, 1987, *Television Graphics: from pencil to pixel*, Trefoil Publications, London

Murray, Janet H, 1997, *Hamlet on the Holodeck: the future of narrative in cyberspace*, The Free Press (Simon and Schuster), New York

Musser, Charles, 1991, *Before the Nickelodeon – Edwin S Porter and the Edison Manufacturing Company*, University of California Press, Berkeley

Nardi, Bonnie A and Zarmer, Craig L, 1993, 'Beyond Models and Metaphors: visual formalisms in user interface design', *Journal of Visual Languages and Computing*, 1993, No 4, p5-33

Naremore, James, 1978, *The Magic World of Orson Welles*, Oxford University Press, New York

Nielsen, Jakob, 2000, www, 'Flash: Ninety-nine Percent Bad', *Alertbox*, October 29, 2000, http://www.useit.com/alertbox/20001029.html (14 November 2000)

Norman, Donald A, 1988, *The psychology of everyday things*, Basic Books, New York

Norman, Donald A, 1999, 'Affordance, Conventions and Design', *Interactions*, Vol.VI, No.3, May/June 1999, p38-42

Oviatt, Sharon and Cohen, Philip, 2000, 'Perceptual user interfaces: multimodal interfaces that process what comes naturally', *Communications of the ACM* (March 2000), Vol.43, No.3, ACM, New York, p45-53

Oxenaar, Rudolf WD, 1982, 'Ven der Leck and De Stijl 1916-1920', in Jaffé, Hans LC (ed.) 1982 *De Stijl: 1917-1931 – Visions of Utopia*, Phaidon, Oxford, p68-79

Panofsky, Erwin, 1991 (1925), *Perspective as Symbolic Form*, Zone Books, New York

Park, David, 1997, *The Fire within the Eye: a historical essay on the nature and meaning of light*, Princeton University Press, Princeton NJ

Parsaye, Kamram and Chignell, Mark, 1993, *Intelligent Database Tools and Applications*, John Wiley and Sons Inc

Pascoe, David, 1997, *Peter Greenaway: museums and moving images*, Reaktion Books, London

Pedersen, Elin Rønby, 1998, 'People Presence or Room Activity Supporting Peripheral Awareness over Distance', *Summary Proceedings of CHI98* Los Angeles 18-23 April 1998, p283-284

Persson, Per, 1998, 'A comparative study of digital and cinematic space with special focus on navigational issues', *Proceedings of Ninth European Conference on Cognitive Ergonomics*, University of Limerick, Ireland, August 1998., p67-72

Pickles, John, 1992, 'Texts, Hermeneutics and Propaganda Maps', in Barnes, Trevor J and Duncan, James S (eds.) *Writing Worlds: discourse, text and metaphor in the representation of landscape*, Routledge, London, p193-230

Pinker, Steven, 1994, *The Language Instinct*, Penguin, London

Podro, Michael, 1998, *Depiction*, Yale University, New Haven

Poole, Steven, 2000, *Trigger Happy: the inner life of videogames*, Fourth Estate, London

Porter, Tom and Susman, Galyn, 2000, 'Creating Lifelike Characters in Pixar Movies', *Communications of the ACM*, January 2000, Vol.43, No.1, ACM, New York, pp25-29

Postman, Neil, 1987, *Amusing Ourselves to Death: public discourse in the age of show business*, Methuen, London

Preece, Jenny; Rogers, Yvonne; Sharp, Helen; Benyon, David; Holland, Simon and Carey, Tom, 1994, *Human Computer Interaction*, Addison Wesley, Harlow, England

Reeves, B and Nass C, 1998, *The Media Equation*, Cambridge University Press

Reisz, Karel and Millar, Gavin, 1982 (1968), 2nd edn., *The Technique of Film Editing*, Focal Press, London

Renoir, Jean, 1974, *My Life and my Films* (translated from the French by Norman Denny), Collins, London

Richards, Clive James, 1984, *Diagrammatics: an investigation aimed at providing a theoretical framework for studying diagrams and for establishing a taxonomy of their fundamental modes of graphic representation*, unpublished PhD thesis, Royal College of Art, London

Richards, Clive James, 2000, 'Getting the Picture: diagram design and the information revolution', *Information Design Journal*, Vol 9, Nos. 2&3, p87-100

Rimmon-Kenan, Shlomith, 1983, *Narrative Fiction: contemporary poetics*, Methuen, London

Rose, Gillian; **Gregson, Nicky**; **Foord, Jo**; **Bowlby, Sophie**; **Dwyer, Claire**; **Holloway, Sarah**; **Laurie, Nina**; **Maddrell, Avril and Skelton, Tracy,** 1997, Introduction, in Women and Geography Study Group (eds.) *Feminist Geographies: explorations in diversity and difference*, Addison Wesley Longman, Harlow, England

Rosenbloom, Andrew, 1999, 'Towards an Image Indistinguishable from Reality', *Communications of the ACM*, August 1999, Vol. 42, No. 8, ACM, New York, p28-30

Sabin, Roger, 1996, *Comics, Comix and Graphic Novels a history of comic art*, Phaidon, London

Salt, Barry, 1990, 'Film Form 1900-1906', in *Early Cinema – Space, Frame, Narrative* edited by Thomas Elsaesser and Adam Barker, BFI Publishing, London, p31-44

Scheepers, Ferdi; **Parent, Richard E**; **Carlson, Wayne E and May, Stephen F,** 1997, 'Anatomy-based modeling of the human musculature', *Proceedings of the 24th annual conference on Computer graphics & interactive techniques*, 1997, ACM, New York, p163-172

Schivelbusch, W, 1978, 'Railroad Space and Railroad Time', *New German Critique*, 14, pp31-40

Schöffel, Frank, 1997, 'Online radiosity in interactive virtual reality applications', *Proceedings of the ACM symposium on Virtual reality software and technology*, 1997, Lausanne Switzerland, ACM, New York, p201-208

Schofield, Simon, 1996, 'Piranesi, a 3-D paint system', *Proceedings of Eurographics UK Annual Conference 1996* (Vol 2), edited by H Jones, R Raby and D Vicars, Imperial College, London, 26-28 March 1996, p91-100

Scrivener, Stephen AR and Clark, Sean M, 1994, 'Sketching in Collaborative Design', in MacDonald, Lindsay and Vince, John (eds.) *Interacting with Virtual Environments*, John Wiley and Sons, Chichester

Searle, John R, 1969, *Speech Acts: an essay in the philosophy of language*, Cambridge University Press, Cambridge

Searle, John R, 1980, 'Las Meninas and the Paradoxes of Pictorial Representation', in Mitchell, WJT (ed.) *The Language of Images*, University of Chicago Press, Chicago, p247-258

Shneiderman, Ben, 1992, 2nd edn., *Designing the User Interface – strategies for effective human-computer interaction*, Addison Wesley, Reading, MA

Smith, Dan, 1999, 6th edn., *The State of the World Atlas*, Penguin, London

Snyder, Joel, 1980, 'Picturing Vision', in Mitchell, WJT (ed.) *The Language of Images*, University of Chicago Press, Chicago, p219-246

Soler, Cyril and Sillion, François X, 1998, 'Fast calculation of soft shadow textures using convolution', *Proceedings of the 25th annual conference on Computer Graphics*, 1998, ACM, New York, p321-332

Solso, Robert L, 1994, *Cognition and the Visual Arts*, MIT, Cambridge MA

Sontag, Susan, 1977, *On Photography*, Penguin (Allen Lane), London

Standage, Tom, 1998, *The Victorian Internet*, Weidenfeld and Nicholson, London

Sutcliffe, Thomas, 2000, *Watching*, Faber and Faber, London

Taussig, Michael, 1993, *Mimesis and Alterity: a particular history of the senses*, Routledge, New York and London

Taylor, Brandon, 1987, *Modernism, Postmodernism, Realism*, Winchester Shcool of Art Press, Hampshire, UK

Taylor, Mark C and Saarinen, Esa, 1994, 'Speed' in *Imagologies – Media Philosophy*, Routledge, London, Chapter 17, no page numbers

Thrift, Nigel, 1996, *Spatial Formations*, Sage Publications, London

Thwaites, Hal, 1999, 'Visual Design in Three Dimensions', Jacobson, Robert (ed.) *Information Design*, MIT Press, Cambridge MA, p221-246

Toy Story, 2000, *Inside Film* (promotional leaflet for Toy Story 2), Studio Seven, London

Tufte, Edward R, 1983, *The Visual Display of Quantitative Information*, Graphics Press, Cheshire, Connecticut

Tufte, Edward R, 1990, *Envisioning Information*, Graphics Press, Cheshire, Connecticut

Tufte, Edward R, 1997, *Visual Explanations: images and quantities, evidence and narrative*, Graphics Press, Cheshire, Connecticut

Twyman, Michael, 1986, 'Articulating Graphic Language: a historical perspective', in Wrolstad, Merald E and Fisher, Dennis F, *Towards a New Understanding of Literacy*, Praeger, New York, 188-251

Usoh, Martin; Arthur, Kevin; Whitton, Mary C; Bastos, Rui; Steed, Anthony; Slater, Mel and Brooks, Frederick P, Jr., 1999, 'Walking > Walking-in-Place > Flying', in *Virtual Environments*, SIGGRAPH 99, Los Angeles, CA USA, ACM, New York, p359-364

Vaz, Mark Cotta and Hata, Shinji, 1995, *The Star Wars Archives*, Virgin Publishing, London

Vince, John, 1995, *Virtual Reality Systems*, Addison Wesley, Wokingham, UK

Virilio, Paul, 1994 (1988), *The Vision Machine* (trans. by Julie Rose from La Machine de Vision, Editions Galilée, Paris 1988), British Film Institute, London

Wainer, Howard, 1997, *Visual Revelations: graphic tales of fate and deception from Napoleon Bonaparte to Ross Perot*, Copernicus / Springer-Verlag, New York

Waters, Keith, 1987, 'A muscle model for animating three-dimensional facial expression', *Computer Graphics (SIGGRAPH 87 Conference Proceedings)*, July 1987, ACM, New York, p17–24

Weinhaus, Frederick M and Devarajan, Venkat, 1997, 'Texture Mapping 3D Models of Real-World Scenes', *ACM Computing Surveys*, December 1997, Vol. 29, No. 4, ACM, New York, p325-365

Willats, John, 1990, 'The Draughtsman's Contract: creating an image', in Barlow, Horace; Blakemore, Colin and Weston-Smith, Miranda (eds) *Images and Understanding*, Cambridge University Press, Cambridge, p235-254

Williams, Frederick S, 1852, *Our Iron Roads – their history, construction and social influences*, Ingram, Cooke and Co, London

Wollheim, Richard, 1980, 2nd edn, *Art and its Objects*, Cambridge University Press, Cambridge UK

Digital products and projects

Art of Invention, 1995, *Instrumentor*, Produced and designed by Art of Invention, Brighton, UK

Chung, Gumo; Fukner, Kirsten; Hoffman, Hans and Rousselot, Nathalie, 2000, *Virtual Office*, Postgraduate Project, Lansdown Centre for Electronic Arts, Middlesex University, UK

Codemasters, 2000, *MicroManiacs*, Playstation game, Codemasters, UK

Cyan Incorporated, 1993, Version 1.2, CD-ROM: *Cosmic Osmo and the Worlds Beyond the Mackerel*, Design by Robyn Miller and Rand Miller; Produced by Cyan Inc; Published by Broderbund Inc.

Dawkins, 1996, CD-ROM: *The Evolution of Life with Richard Dawkins*, Editorial Content by Richard Dawkins, London

Dorling Kindersley, 1997, Sampler of various CD-ROM titles, on Macaulay 1994, Produced and published by Dorling Kindersley, London

Eberle, Lars, 1996, Digital animation: *Upholstery Weekend*, Postgraduate Project, Lansdown Centre for Electronic Arts, Middlesex University, UK

Editoriale Domus, 1998, CD-ROM: *Venezia*, In series le Città d'Arte, Editoriale Domus. Project designed and realised by Gabo Multimedia SnC, Venice, Italy

Foreign and Commonwealth Office, 1997, CD-ROM: *Britain in Brief*, Design and production by Art of Invention and The Central Office of Information, London

Good Technology, 2000, Website: U2, http://www.goodtechnology.com (16 November 2000)

Holley, Tom; Reeves, John; Sauderais, Magali; Sjaastad, Stine; Choy, Kok Kee, 1998, *Eclipse: a prototype interactive fiction*, MA Masters Project at Lansdown Centre for Electronic Arts at Middlesex University, UK, unpublished prototype

Lac, Visieu; Wölwer, Stefan and Wu, Mark, 1998, Interactive demonstration: *Esfore-Entropy*, Postgraduate Project, Lansdown Centre for Electronic Arts, Middlesex University, UK

Maltez Dulce, Bennett, Brett and Cova, Mattia, 1997, *Contact: a prototype interactive documentary*, MA Masters Project at Lansdown Centre for Electronic Arts at Middlesex University, UK, unpublished prototype

Miller, Rand and Robin, 1993, CD-ROM: *Myst*, Produced by Cyan. Published by Broderbund

Norman, Donald A, 1994, CD-ROM: *Donald A Norman – defending human attributes in the age of the machine*, Produced and published by Voyager, USA

Notting Hill/JHM, 1996, CD-ROM: *The Art of Singing*, Editorial Content: Wilf Judd & Felicity Hayes-McCoy; Art Director: Tim Warren; Published by Notting Hill Publishers / JHM, London

Open University, 1997, CD-ROM: *M206 Interactive Course Map*, Distributed to enrolled students by Open University; Produced by Neil Edwards, Rob Griffiths, Diane Mole, Mark Rowe, David Saunders, Tamara Sumner, Josie Taylor and David Winter, Milton Keynes, UK

PopTop Software, 1998, CD-ROM Strategy Game: *Railroad Tycoon II*, Published by Gathering of Developers Inc, USA and Take 2 Interactive Software, Berkshire, UK

Sim-Business, 1993, Floppy-disc-based strategy game: *SimCity 2000*, Published by Maxis, London; Distributed by Mindscape, Burgess Hill, West Sussex, UK

V&A/BBC, 2000, CD-ROM: *A History of Britain: treasures from the Victoria and Albert museum which link to Simon Schama's acclaimed BBC Television series*, Published by BBC History Magazine; production by Atticmedia Northwest, UK

Films

Cameron, James, 1997, *Titanic*, TCF, Lightstorm, US, 194 minutes DeLuxe Panavision

Chabrol, Claude, 1968, *La Femme Infidèle*, La Boëtie / Cinegay, France / Italy, 98 minutes Eastmancolor

Curtiz, Michael, 1942, *Casablanca*, Warner, US, 102 minutes b / w

Fleming, Victor (with George Cukor and Sam Wood), 1939, *Gone with the Wind*, Selznick International / Metro Goldwyn Meyer, US, 220 minutes Technicolor

Gance, Abel, 1927, *Napoleon*, WESTI / Société Générale des Films, France, 378 minutes approx (@24fps) b / w with some colour tinting

Greenaway, Peter, 1982, *The Draughtsman's Contract*, BFI / Channel Four, GB, 108 minutes colour

Greenaway, Peter, 1985, *Zed and Two Noughts*, Artificial Eye, BFI, Allarts Enterprises, Film Four International, GB, 115 mintues colour

Greenaway, Peter, 1991, *Prospero's Books*, Allarts-Cinéa, Camera One-Penta with Elsevier Vendex Film, Film Four International, NHK Enterprises, Canal + and VPRO Television, Netherlands / France / Italy, 120 minutes colour

Hitchcock, Alfred, 1941, *Suspicion*, RKO, US, 99 minutes b / w

Lasseter, John, 1995, *Toy Story*, Walt Disney / Buena Vista / Pixar, US, 80minutes Technicolor

Lasseter, John, 2000, *Toy Story 2*, Walt Disney / Buena Vista / Pixar, US, 88minutes Technicolor

Lean, David, 1962, *Lawrence of Arabia*, Columbia / Horizon, GB, 221 minutes Technicolor, Super Panavision

Montgomery, Robert, 1946, *The Lady in the Lake*, MGM, US, 103minutes b / w

Nicholls, Mike, 1966, *Who's Afraid of Virginia Woolf?*, Warner, US, 129 minutes b / w

Polanski, Roman, 1968, *Rosemary's Baby*, Paramount, US, 137minutes Technicolor

Powell, Michael, 1959, *Peeping Tom*, Anglo Amalgamated, GB, 109 minutes Eastmancolor

Reed, Carol, 1949, *The Third Man*, British Lion, London Films, David O Selznick, Alexander Korda, GB, 100 minutes b / w

Scorsese, Martin, 1990, *Goodfellas*, Warner, US, 146 minutes Technicolor

Welles, Orson, 1941, *Citizen Kane*, RKO, US, 119 minutes b / w

Welles, Orson, 1958, *Touch of Evil*, U-I, US, 95 or 114 minutes b / w